

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ
КАФЕДРА МАТЕМАТИЧНИХ МЕТОДІВ ЗАХИСТУ ІНФОРМАЦІЇ

«На правах рукопису»
УДК _____

«До захисту допущено»

В.о. завідувача кафедрою
_____ М.М.Савчук
(підпис) (ініціали, прізвище)

“ ____ ” _____ 2020р.

Магістерська дисертація

на здобуття ступеня магістра

зі спеціальності Прикладна Математика _____
(код і назва)

на тему: Збільшення просторового розрізнення температурних продуктів
Sentinel-3 з використанням рекурентних нейронних мереж _____

Виконав студент _____ 6 курсу, групи ФІ-83мн _____
(шифр групи)

(прізвище, ім'я, по батькові) _____ (підпис)

Керівник Д. Фізико-Математичних наук Савчук М.М. _____
(посада, науковий ступінь, вчене звання, прізвище та ініціали) _____ (підпис)

Консультант 1-5 _____ Д. Технічних наук Куссуль Н.М. _____
(назва розділу) (науковий ступінь, вчене звання, прізвище, ініціали) _____ (підпис)

Рецензент _____
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали) _____ (підпис)

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць
інших авторів без відповідних
посилань.

Студент _____
(підпис)

Київ – 2020_року

Національний технічний університет України
«Київський політехнічний інститут
імені Ігоря Сікорського»
Фізико-технічний інститут
Кафедра математичних методів захисту інформації

Рівень вищої освіти: другий (магістерський) за освітньо–науковою програмою

Спеціальність: 113 «Прикладна математика»

ЗАТВЕРДЖУЮ

В.о. завідувача кафедрою

_____ М.М.Савчук
(підпис) (ініціали, прізвище)

«___» _____ 20_ р.

ЗАВДАННЯ

на магістерську дисертацію студенту

Шуміла Леоніда Леонідовича _____

(прізвище, ім'я, по батькові)

1. Тема дисертації: Збільшення просторового розрізнення температурних продуктів Sentinel-3 з використанням рекурентних нейронних мереж, науковий керівник дисертації Д. Фізико-Математичних наук Савчук М.М. _____,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від _____ р. № _____

2. Термін подання студентом дисертації _____

3. Об'єкт дослідження є карти температури земної поверхні отримані на основі супутникових даних Sentinel-3 та моделі штучної нейронної мережі LSTM.

4. Предмет дослідження (Вхідні дані – для магістерської дисертації за освітньо–професійною програмою)

Дані супутникових місій Landsat-8, Sentinel-2, Sentinel-3

5. Перелік завдань, які потрібно розробити Провести аналіз існуючих способів злиття даних для збільшення просторового розрізнення супутникових даних, та збільшення просторового розрізнення супутникових даних; Покращити метод відновлення хмар на супутникових продукті LST та NDVI; Розробити математичний метод глибинного навчання для збільшення просторового розрізнення супутникових даних температури земної поверхні; Підготувати навчальний та тестовий набори даних, а також провести його статистичний аналіз; Використати супутникові продукти температури земної

поверхні для прикладних досліджень за напрямком технологій розумних міст.

6. Орієнтовний перелік ілюстративного матеріалу наведено 45 ілюстрацій, з яких 20 ілюструють оброблені та необроблені дані, 6 процеси обробки даних та 19 отримані результати

7. Орієнтовний перелік публікацій в даній роботі наведено 70 публікацій серед яких вказано публікації створені під час роботи над дисертацією, посилання на використані методи обробки даних та аналізу даних, а також на роботи в яких використовуються ці дані та методи

8. Консультанти розділів дисертації*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1-5	Д. Технічних наук Куссуль Н.М.		

9. Дата видачі завдання 2 вересня 2018

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Аналіз джерел супутникових даних	Вересень 2018 – Жовтень 2018	
2	Побудова досліджень для міст України на основі супутникових даних	Листопад 2018 – Лютий 2019	
3	Імплементация методології виявлення пожеж на основі даних Sentinel-3	Лютий 2019 – Липень 2019	
4	Підготовка та гармонізація супутникових даних iSentinel-2, Landsat-8, Sentinel-3	Липень 2019 – Листопад 2019	
5	Розробка методу збільшення просторового розрізнення супутникових даних Sentinel-3	Листопад 2019 – Лютий 2020	
6	Тестування розробленої методології	Лютий 2020 – Квітень 2020	

Студент

_____ (підпис)

_____ (ініціали, прізвище)

Науковий керівник дисертації

_____ (підпис)

_____ (ініціали, прізвище)

* Консультантом не може бути зазначено наукового керівника магістерської дисертації.

РЕФЕРАТ

Кваліфікаційна робота містить: 100 стор., 45 рисунків, 5 таблиць, 70 джерел.

Температура земної поверхні (LST) є важливим показником стану міських територій та може використовуватися для виявлення проблемних зон міста та моніторингу природних лих. Супутникові дані є єдиним джерелом інформації, що дає можливість проводити аналіз LST для великих територій. Для цього використовуються алгоритми моделювання LST на основі термальних каналів мультиспектральних супутникових даних.

В рамках проекту ERA-Planet SMURBS Інститут Космічних Досліджень НАНУ та ДКАУ розробляє сервіси моніторингу та аналізу розумних міст. Важливими сервісами є моніторинг температури земної поверхні міста, аналіз планування міста за допомогою характеристик поверхні землі та виявлення підземних та наземних пожеж. Двома основними актуальними місіями, які надають мультиспектральні дані для моделювання температури земної поверхні є місія Sentinel-3 та місія Landsat-8. Проте різне просторове та часове розрізнення спричиняє складнощі у поєднанні даних отриманих від цих супутникових місій. Landsat-8 надає супутникові дані, що дозволяють отримувати карти температури земної поверхні з високим просторовим розрізненням термального каналу 30 м., проте має низьке часове розрізнення 16 днів, що є недостатнім для сервісів моніторингу температури земної поверхні в режимі близькому до реального часу. Місія Sentinel-3 дає можливість отримувати карти температури земної поверхні з просторовим розрізнення 1 км., що робить ці знімки мало інформативними для аналізу міської поверхні, при цьому ця місія має високе часове розрізнення в 2 знімки на день (ранковий та вечірній час). Враховуючи ці специфіки даних виникає актуальна задача збільшення просторового розрізнення

LST продуктів Sentinel-3. Це є можливим при застосуванні сучасних методів злиття мультиспектральних супутникових даних та методів глибинного навчання.

Результатом даної дисертації є новий метод збільшення просторового розрізнення супутникових даних Sentinel-3 на основі злиття даних Sentinel-2, Landsat-8 та Sentinel-3 і архітектури рекурентної нейронної мережі LSTM. Покращено метод відновлення хмар супутникових продуктів на основі регресійної моделі випадкових лісів, яка застосовується для виявлення не замаскованих пікселів хмар в якості аномалій і відновлення хмар. Зроблено прикладні дослідження з використанням супутникових даних для міста Рівне, Ірпінь та Київ.

ТЕМПЕРАТУРА ЗЕМНОЇ ПОВЕРХНІ, НЕЙРОННІ МЕРЕЖІ,
ДИСТАНЦІЙНЕ ЗОНДУВАННЯ ЗЕМЛІ, LONG SHORT TERM
MEMORY RECURRENT NEURAL NETWORKS, SENTINEL-3,
РОЗУМНІ МІСТА

ABSTRACT

Land Surface Temperature (LST) is an essential urban variable and can be used to identify problem areas of the city and monitor of natural disasters. Satellite data is the only source of information that give possibility provide LST analysis in large areas. For this purpose, LST modeling algorithms based on thermal channels of multispectral images are used.

Within ERA-Planet SMURBS project Space Research Institute NASU-SSAU developing smart cities monitoring and analysis services. These services include smart cities land surface temperature monitoring, analysis of the city planning using the land surface characteristics and peat-fires and active fire detection. The two major satellite missions that provide multispectral data for simulating land surface temperature are the Sentinel-3 mission and the Landsat-8 mission. However, different spatial and temporal resolution make it difficult to combine the data obtained from these satellite missions. The Landsat-8 provides satellite data that allows to obtain land surface temperature maps with a high spatial resolution of the 30 m , but has a low temporal resolution of 16 days, which is not sufficient for near-real-time operational land surface temperature monitoring services. The Sentinel-3 mission provides 1-km-wide land surface temperature maps and this fact makes these images less informative for urban surface analysis, with a high temporal resolution of 2 images per day (morning and evening). Based on these data specifics, an important task arises to increase the LST spatial resolution of the Sentinel-3 products. This is possible by using modern methods of multispectral satellite data fusion and deep learning methods.

The result of this dissertation is a new method of Sentinel-3 satellite data spatial resolution up-scaling based on the fusion of the Sentinel-2, Landsat-8 and Sentinel-3 data and the LSTM recurrent neural network architecture. Improved satellite product cloud recovery method based on random forest regression model used to detect non-masked pixel clouds as

anomalies and cloud recovery. Applied research has been done using satellite⁷
data for the cities of Rivne, Irpin and Kyiv.

LAND SURFACE TEMPERATURE, NEURAL NETWORK,
REMOTE SENSING, LONG SHORT TERM MEMORY RECURRENT
NEURAL NETWORKS, SENTINEL-3, SMART CITY

ЗМІСТ

Перелік умовних позначень, скорочень і термінів	8
Вступ.....	9
1 Задача збільшення просторового розрізнення супутникових продуктів	13
1.1 Відкриті глобальні супутникові місії	13
1.2 Класи просторового розрізнення	18
1.3 Пан-шарпенінг	19
1.4 Міський Атлас, як векторний шар для геопросторовго аналізу ...	21
Висновки до розділу 1	23
2 Моделювання температури земної поверхні на основі супутникових даних	25
2.1 Основні підходи моделювання темперартури земної поверхні на основі супутникових даних	25
2.2 Split Window Algorithm	26
2.3 Radiative Transfer Equasion	27
Висновки до розділу 2	31
3 Нейронно мережева методологія збільшення просторового розрізнення температурних продуктів Sentinel-3	32
3.1 Рекурентні нейронні мережі	32
3.2 Архітектура нейронної мережі з довгою та коротко часовою пам'яттю	35
3.3 Використання нейронної мережі Long Short Term Memory для збільшення просторового розрізнення температурних карт	37
Висновки до розділу 3	44
4 Використання методології збільшення просторового розрізнення супутникових карт температури земної поверхні Sentinel-3	45
4.1 Підготовка та гармонізація суптуникових даних	45
4.2 Відновлення хмар на супутникових продуктах	49

4.3 Формування навчальних і тестових даних для моделі збільшення просторового розрізнення супутникових даних	55
4.4 Тестування моделі нейронної мережі для збільшення просторового розрізнення супутникових даних	63
Висновки до розділу 4.....	66
5 Використання супутникових продуктів температури земної поверхні у прикладних задачах	69
5.1 Аналіз характеристик поверхні міста для виявлення проблемних ділянок у місті Рівне.....	70
5.2 Використання супутникових продуктів для розробки плану озеленення міста	77
5.3 Створення рекомендацій для програми озеленення міста Ірпінь ..	79
5.4 Виявлення пожеж за допомогою карт температури земної поверхні	84
Висновки до розділу 5.....	89
Висновки	91
Перелік посилань	93

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

У даній дисертації вживаються такі скорочення:

LST — Температура Земної Поверхні

LSTM — Long Short Term Memory

RGBN — Red, Green, Blue, Nir Red

NDVI — Normalized Difference Vegetation Index

ГІС — Гео Інформаційні Системи

TOA — відбиття на вершині атмосфери

NIR — відбиття в інфрачервоному спектрі

ВСТУП

Актуальність дослідження. Дані дистанційного зондування землі (ДЗЗ) сьогодні мають велику кількість застосувань у різних напрямках людського життя, починаючи від сільського господарства, закінчуючи розвідкою та моніторингом екології. Це стало можливо після запуску програми Copernicus, що спричинило збільшення кількості доступних супутникових даних та готових аналітичних продуктів, що постійно оновлюються. Таким чином з'явилася потреба у нових методах обробки даних та збільшення актуальності досліджень у напрямку використання штучного інтелекту для роботи з супутниковими даними. Новітні архітектури глибоких нейронних мереж вже показали свою ефективність у багатьох прикладних задачах.

В ході реалізації сервісів в проєкті ERA-Planet SMURBS (SMart URBan Solutions for air quality, disasters and city growth) на основі супутникових даних температури земної поверхні для моніторингу та аналізу міст та виявлення пожеж, з'явилася потреба у створенні колекції даних температури земної поверхні з високим просторовим та часовим розрізненням. Для цього необхідно поєднати особливості доступних мультиспектральних супутникових даних, а саме високе просторове розрізнення температурних карт Landsat-8 та високе часове розрізнення супутникових продуктів Sentinel-3. Використовуючи сучасні методи поєднання мультиспектральних супутникових даних із особливостями штучних нейронних мереж архітектури LSTM, дають можливість збільшити просторове розрізнення температурних карт Sentinel-3, таким чином задовольнити потреби сервісів SMURBS. При цьому актуальність розробки таких методів обробки супутникових даних підвищується тим, що супутникова місія Sentinel-3 є новою місією Європейського Космічного Агенства, запущена у 2016-ому році і на сьогодні є велика потреба у розробці нових методів, або модифікації старих для роботи з цими

супутниковими даними.

Метою дослідження є розробка математичного методу на основі глибинного навчання з застосуванням архітектур LSTM для збільшення просторового розрізнення супутникових продуктів температури земної поверхні для отримання даних температури земної поверхні з високим просторовим та часовим розрізненням

задачі дослідження, які полягає у об'єднанні різномірних супутникових та інших геопросторових даних для отримання цифрових продуктів аналізу міста та розробці методології збільшення просторового розрізнення карт температури земної поверхні Sentinel-3 з використанням архітектури рекурентної штучної нейронної мережі типу LSTM. Для розв'язання цієї задачі необхідно вирішити такі завдання:

- 1) Провести аналіз існуючих способів злиття даних для збільшення просторового розрізнення супутникових даних, та збільшення просторового розрізнення супутникових даних;
- 2) Покращити метод відновлення хмар на супутникових продукті LST та NDVI;
- 3) Розробити математичний метод глибинного навчання для збільшення просторового розрізнення супутникових даних температури земної поверхні;
- 4) Підготувати навчальний та тестовий набори даних, а також провести його статистичний аналіз;
- 5) Використати супутникові продукти температури земної поверхні для прикладних досліджень за напрямком технологій розумних міст.

Об'єктами дослідження є карти температури земної поверхні отримані на основі супутникових даних Sentinel-3 та моделі штучної нейронної мережі LSTM.

Предметами дослідження є моделі поєднання мультиспектральних супутникових продуктів та застосування моделей рекурентних нейронних мереж у процесі збільшення просторового розрізнення супутникових продуктів температури земної поверхні Sentinel-3.

При розв'язанні поставлених завдань використовувались такі *методи дослідження*: фізика, лінійна алгебра, штучний інтелект, глибинне навчання, методи обробки даних у хмарному середовищі Google Earth Engine та Amazon EC2.

Наукова новизна полягає у новій методології збільшення просторового розрізнення карт температури земної поверхні на основі супутникових даних Sentinel-3 та імплементація цієї методології у сервіси моніторингу розвитку міста та природніх лих проєкту ERA-Planet SMURBS.

Практичне значення полягає у використанні отриманих результатів у розробці сервісів розумного міста проєкту ERA-Planet SMURBS та створенні нової колекції даних температури земної поверхні з високим просторовим та часовим розрізненням.

Апробація результатів та публікації. Результати отримані при виконанні даної дипломної роботи опубліковані у збірниках конференцій:

1) 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON) — Land Surface Temperature Analysis and Trends for Rivne Using Remote Sensing Data

2) 2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT) — Sentinel-3 Urban Heat Island Monitoring and analysis for Kyiv Based on Vector Data

3) 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO) — Rivne City Land Cover and Land Surface Temperature Analysis Using Remote Sensing Data

Презентовано на конференціях:

1) Living Planet Symposium 2019 — Land Surface Temperature estimation for Smart City

Прийняті тези:

1) Internationall Geo science and Remote Sensing Symposium 2020 (IGARSS) — ACTIVE FIRE MONITORING SERVICE FOR UKRAINE BASED ON SATELLITE DATA

2) EO for Agriculture under Pressure 2020 — Sentinel-3 Application For
Stubble Burn Identification In Ukraine¹²

1 ЗАДАЧА ЗБІЛЬШЕННЯ ПРОСТОРОВОГО РОЗРІЗНЕННЯ СУПУТНИКОВИХ ПРОДУКТІВ

Просторове розрізнення, це одна із головних характеристик супутникових продуктів. По суті просторове розрізнення вказує на мінімальний розмір земельної ділянки, яка займає на супутникову знімку один піксель. Зазвичай супутникові місії з високим просторовим розрізненням мають низьке часове розрізнення, а місії що мають високе часове розрізнення часто мають низьке просторове розрізнення.

Збільшення кількості супутникових місій протягом останніх 10 років призвело до значного збільшення кількості застосувань супутникових продуктів. Покращення якості продуктів інтелектуальної обробки супутникових даних відбувається через збільшення просторового розрізнення, покращення існуючих методів і збільшення кількості супутникових даних. При цьому потрібно розуміти, що для деяких задач важливішим є збільшення кількості супутникових даних та покращення методів. Саме з відси і впливає задача збільшення просторового розрізнення супутникових даних, для того щоб збільшити кількість даних необхідних для розв'язання прикладних задач.

1.1 Відкриті глобальні супутникові місії

На сьогодні є три найбільш часті у використанні супутникові місії, які є джерелом інформації для моделювання температури земної поверхні. На даний момент у відкритому доступі знаходяться дані 3 популярних супутникові місій, які містять інформацію LST. Також важливою місією моніторингу землі, що не містить температуру земної поверхні є місія Sentinel-2 яка надає RGBN знімки із 10 метровим просторовим розрізненням. Ці дані є доступними для завантаження на

порталах провайдерів, тобто NASA та ESA, проте також потрібно відзначити, що доступно дві платформи за допомогою яких можна отримати доступ до цих даних, візуалізувати їх через ГІС у браузері та провести операції обробки даних.

Першою платформою є EOBrowser [1]. Ця платформа дає доступ до всіх популярних активних та завершених супутникових місій які надають безкоштовні глобальні супутникові продукти. Портал дає можливість шукати супутникові продукти онлайн, візуалізувати їх, відкривати дані готові для аналізу такі як LST, NDVI та інші індекси, проводити математичні операції над каналами знімку за допомогою Java Script в режимі онлайн, а також завантажувати результати обчислень, дані для аналізу та дані для візуалізації.

Другою платформою є Google Earth Engine [2] яка надає доступ до всіх популярних активних та завершених супутникових місій які надають безкоштовні глобальні супутникові продукти, а також до великої кількості окремих відкритих геопросторових датасетів. Платформа дає такі ж можливості, як і EOBrowser, проте важливою відмінністю платформи є безкоштовний доступ до хмарних обчислювальних ресурсів Google та можливість в режимі онлайн писати і використовувати комплексні скрипти на мові програмування Java Script, з використанням всіх доступних датасетів та вже імплементованими базовими функціями обробки даних для розв'язання геопросторових задач [3, 4, 5]. Деякі супутникові дані із сховищ Google можна завантажувати напряму, а також разом з будь-яким результатом обробки даних в Google Earth Engine можна експортувати в особисте хмарне сховище Google або на Google Drive.

Першою супутниковою місією є MODIS, запущена NASA. Супутники MODIS Aqua та Terra мають 250 м просторове розрізнення RGBN знімків та дані температури земної поверхні з просторовим розрізненням 1 км. Часове розрізнення дорівнює 2 знімкам в день. Ця місія почалася у 1999 році і її час життя вже закінчився, проте супутники досі знаходяться на своїх

орбітах та надають на постійній основі супутникові дані. Через відсутність перспективи викорисатння даних цієї місії, немає сенсу її використовувати для створення сервісів моніторингу розумних міст, проте ці оскільки архів цієї місії містить дані за велику кількість років її цінність не зменшується. На рисунку 1.1 показано покриття України знімками MODIS за 31-08-2019.

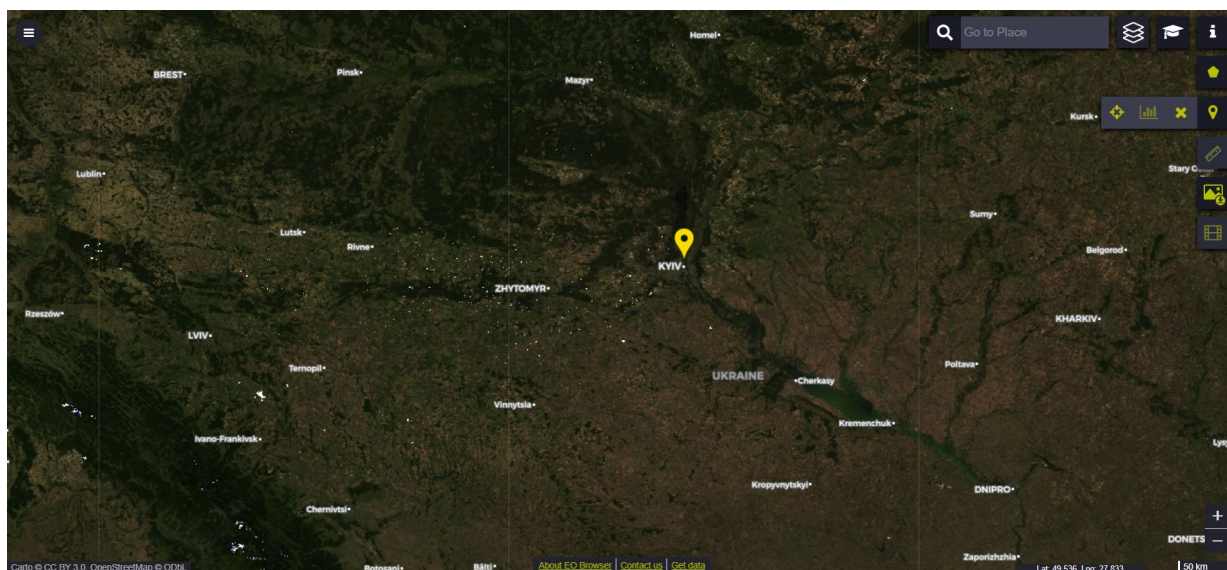


Рисунок 1.1 – Покриття України даними MODIS

Другою супутниковою місією є Landsat-8, запущена NASA. Landsat-8 це супутник з просторовим розрізненням RGBN знімків 30 метрів (рис. 1.2), який також містить панхроматичний канал з розрізненням 15 м, який може бути використаним для збільшення розрізнення до 15 м. Часове розрізнення даного супутника є 16 днів. Термальний канал даного супутника має просторове розрізнення 100 м, проте NASA, як провайдер даних інтерполює його до 30 м та надає ці дані в такому вигляді, тому можна вважати, що просторове розрізнення термального каналу та відповідно карт температури земної поверхні дорівнює 30 м. Цей супутник по просторовому розрізненню є дуже інформативним для проведення аналізу температури земної поверхні, проте часове розрізнення є досить низьким, особливо враховуючи, що це

оптичні дані, які можуть містити хмари і тому кількість доступних знімків які містять корисну інформацію може бути малою.

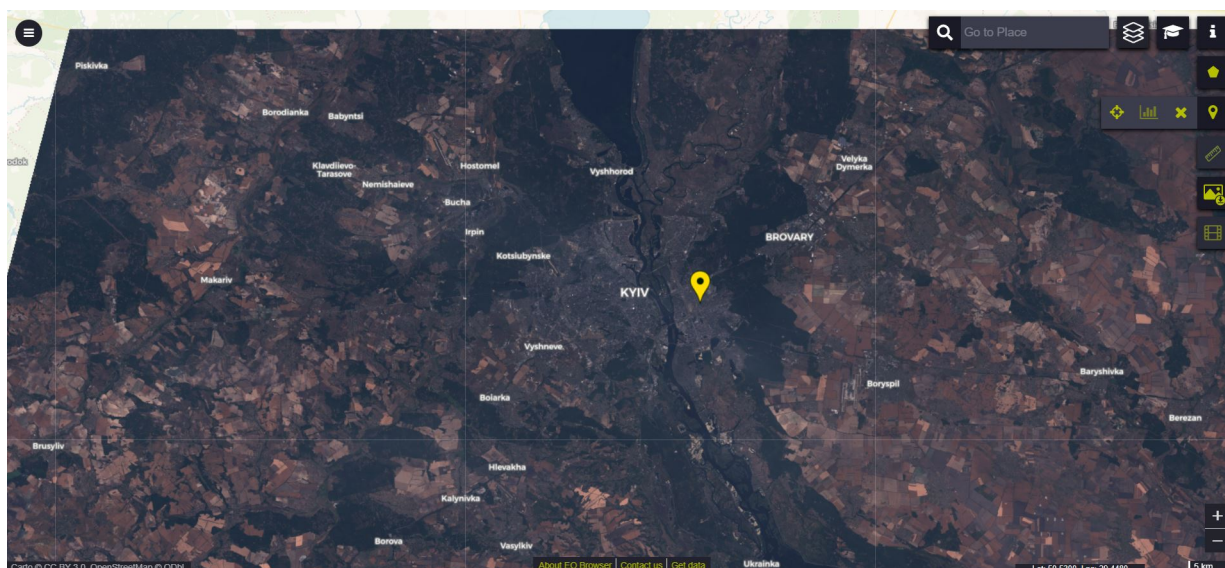


Рисунок 1.2 – Знімок Landsat-8, True Color, що покриває місто Київ

Третьою супутниковою місією є Sentinel-3, запущена ESA. Sentine-3 А та В це пара супутників які мають просторове розрізнення RGBN (рис. 1.3) знімків 300 м та розрізнення термальних знімків 1 км (Рис. 1.4). Часове розрізнення для цих супутників є 2 знімки на день. При цьому заплановано запуск ще двох супутників у цій місії, що підвищить часове розрізнення. Оскільки супутникова місія запущена тільки у 2016-ому році, а дані стали доступними тільки з 2017-го року, наразі ці супутники мають велику перспективу у використанні цих даних для сервісів розумного міста, при умові збільшення розрізнення термальних карт за допомогою злиття даних з інших мультиспектральних джерел. Також важливо зауважити, що температурні карти змодельовані за допомогою split-window algorithm надаються Європейським Космічним Агентством як продукти другого рівня обробки і стають доступними після 30 хв - 1 год. після зйомки.

Місія Sentinel-2 це супутникова місія, яка була запущена Європейським Космічним Агентством в 2015-ому році. Ця місія станом на

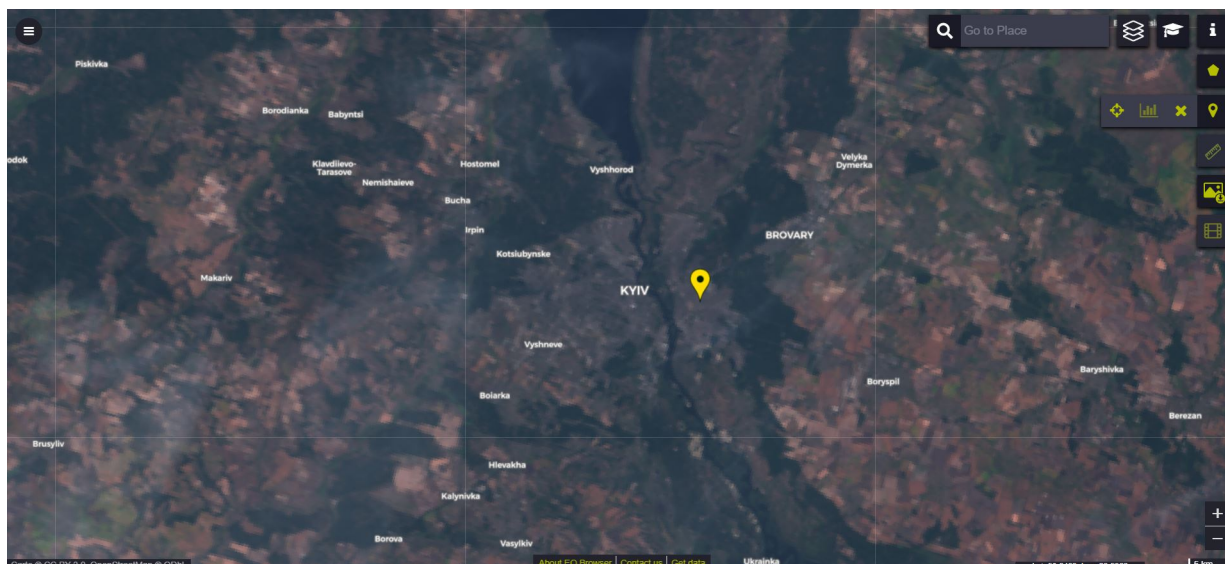


Рисунок 1.3 – Знімки Sentinel-3, True Color, що покриває місто Київ

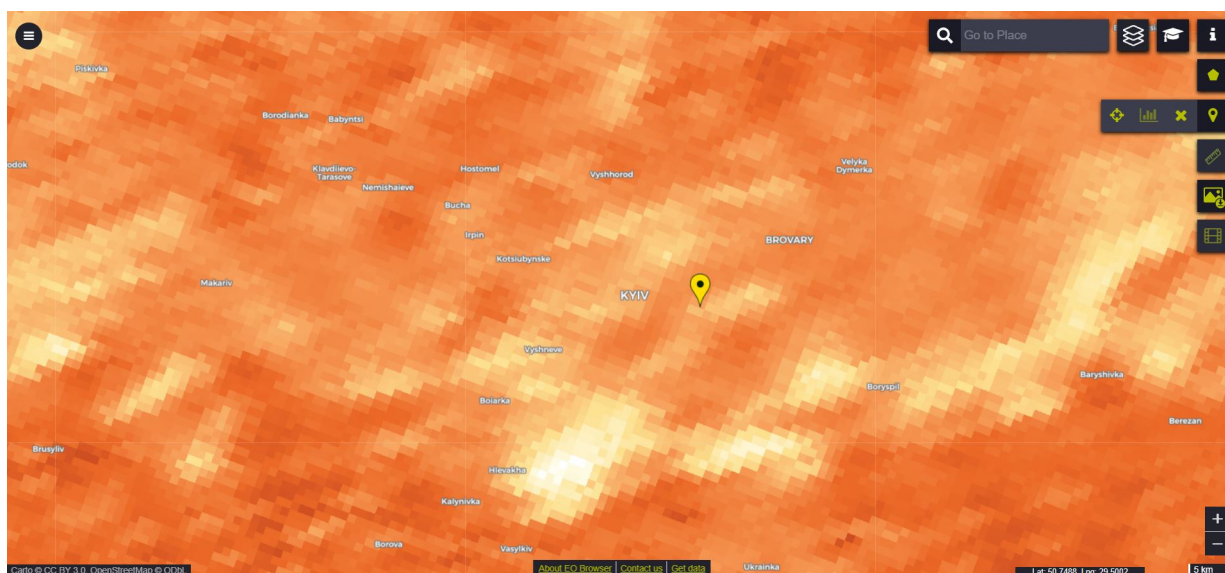


Рисунок 1.4 – Карта температури земної поверхні Sentinel-3 для міста Київ

2020-ий рік складається із двох супутників Sentinel-2A та Sentinel-2B, які надають оптичні знімки RGBN (рис. 1.5) з просторовим розрізненням 10 метрів та часовим розрізненням 5 днів. Оптичні дані цієї місії вирішують багато прикладних задач у сфері моніторингу землі та океанів, особливо вони ефективно себе показують у моніторингу сільського господарства [6, 7, 8] та моніторингу урбанізації [9].

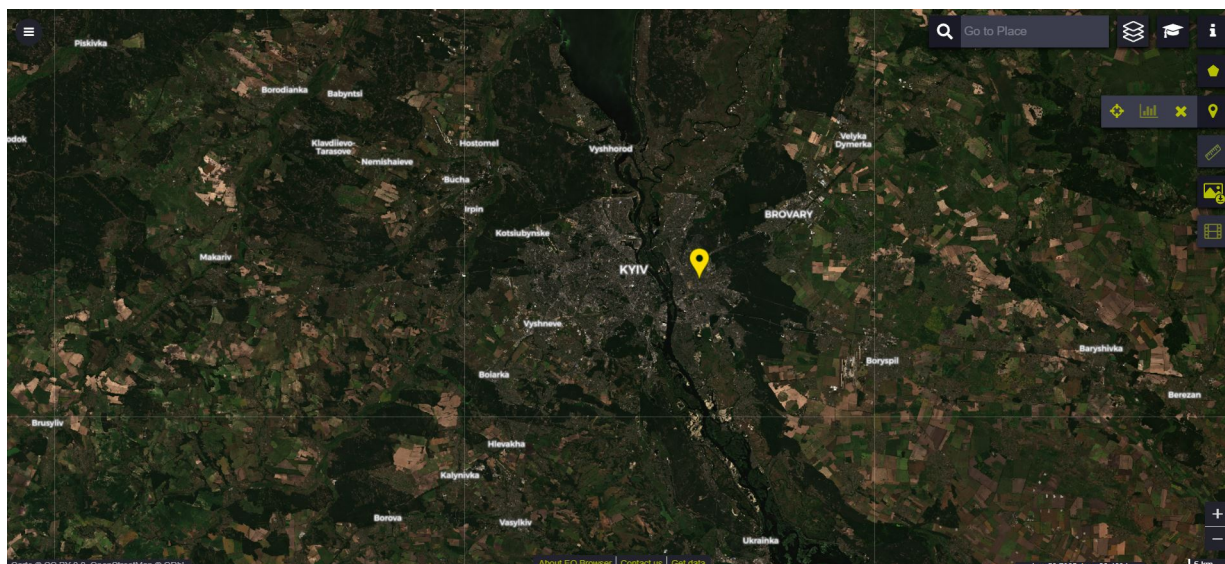


Рисунок 1.5 – Знімок Sentinel-2 True Color, що покриває місто Київ

1.2 Класи просторового розрізнення

Просторове розрізнення це характеристика супутникових даних, яка інтерпретується як розмір поверхні землі, що покривається одним пікселем супутникового знімку. Просторове розрізнення в залежності від розміру одного пікселя можна розділити на 4 класи [10]: низьке, середнє, високе та дуже високе. У таблиці 1.1 наведено діапазони просторового розрізнення та приклади супутникових місій з таким розрізненням.

При цьому варто зауважити, що при розвитку дистанційного зондування землі, діапазони для різних класів просторового розрізнення змінюється. При цьому на законодавчому рівні заборонено використовувати у приватному секторі супутники з роздільною здатністю вище 30 см. При цьому для різних задач використовується дані різного просторового розрізнення. Для моніторингу сільськогосподарських полів краще використовувати дуже високе просторове розрізнення [11], яке дає можливість дуже точно визначити стан насаджень в різних частинах поля. Для моніторингу посівів на рівні країни для України є достатнім високе просторове розрізнення [12]. Для моніторингу сільського господарства на глобальному рівні достатньо середнього просторового

Таблиця 1.1 – Класифікація просторового розрізнення супутникових даних

Назва	Розрізнення (м)	Супутникові місії
Низьке	<1000	Polder
Середнє	100-1000	MODIS AVHRR Sentinel-3
Високе	5-100	Aster Landsat Sentinel-2 Sentinel-1 SPOT
Дуже високе	5<	Quickbird WorldView Geoeye

розрізнення [13]. Таким чином значення класів просторового розрізнення може також залежити від задачі в якій ці дані використовуються.

1.3 Пан-шарпенінг

Пан-шарпенінг це метод об'єднання мультиспектральних супутникових каналів (MS) та панхроматичного (PAN) каналу для збільшення просторового розрізнення мультиспектральних каналів до просторового розрізнення панхроматичного каналу. Для багатьох супутникових оптичних сенсорів високого просторового розрізнення це звичайна практика і тому в багатьох описах продуктів, таких як наприклад WorldView-2, вказується просторове розрізнення для мультиспектральних даних рівне панхроматичному, оскільки для багатьох провайдерів даних паш-шарпенінг це одна із основних процедур обробки даних.

Серед глобальних супутникових місій, дані яких надаються у

відкритому доступі хорошим прикладом є місія Landsat-8. Супутникові знімки Landsat-8 мають просторове розрізнення 30 м, при цьому розрізнення пан-хроматичного каналу дорівнює 15 метрам [14]. На рисунку 1.6 зображено приближення до території національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» міста Київ на даних Landsat-8, зліва RGB знімок, справа панхроматичне зображення.

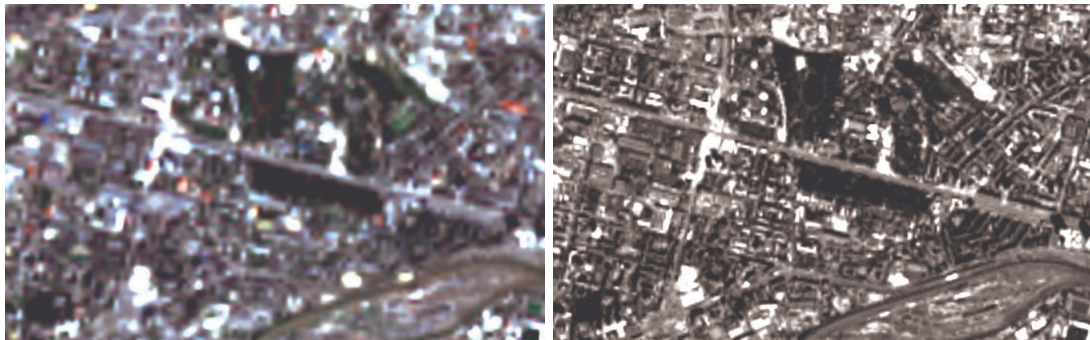


Рисунок 1.6 – 30-ти метровий RGB знімок та 15-ти метровий панхроматичний канал для Києва

Паншарпенінг здійснюється за рахунок об'єднання інформації, яка зберігається у пікселях вищого просторового розрізнення панхроматичного каналу, пікселях нижчого просторового розрізнення мультиспектральних даних та інтенсивності мультиспектрального світла, яка детектується сенсором. Ця інтенсивність (I) вираховується за формулою:

$$\frac{0.6 * Red + 0.98 * Green + 0.42 * Blue}{2}$$

Де Red, Green, Blue це інтенсивність випромінювання зафіксованого у червоному, зеленому та синьому спектральних каналах. Або також при застосуванні NIR каналу інтенсивність вираховується за формулою:

$$\frac{0.6 * Red + 0.98 * Green + 0.42 * Blue + NIR}{3}$$

Таким чином відбувається злиття інформації мультиспектральних

каналів супутникових знімків, в ході якого отримуються знімки вищого просторового розрізнення ніж оригінальні знімки. На рисунку 1.7 зображено порівняння оригінального знімку Landsat-8 та результату злиття. Як видно на знімку з'явилися деталі із панхроматичного каналу. При цьому існує велика кількість різних модифікацій методів пан-шарпенінгу [15, 16] , при цьому існує правило - при застосуванні методу паншарпенінгу збільшується просторове розрізнення, при цьому втрачається мультиспектральне розрізнення і навпаки. Не існує ідеального методу, який одночасно збільшує просторове розрізнення і мультиспектральне розрізнення. В результаті є можливість підвищити інформативність даних для різних задач дистанційного зондування землі [17, 18].



Рисунок 1.7 – 30-ти метровий RGB знімок та 15-ти метровий RGB знімок для Києва

1.4 Міський Атлас, як векторний шар для геопросторового аналізу

Європейською комісією було реалізовано проект Європейський міський атлас (European Urban Atlas) [19], який є частиною локального компонента сервісу моніторингу земель Copernicus Land Monitoring Service (Рис. 1.8.). Він створений на основі використання супутникових

продуктів високого просторового розрізнення на основі яких проводиться класифікація земного покриву. В поєднанні із векторною сегментацією міста цей сервіс надає інтерактивну карту-план міста. Такі карти будуються кожна 4-5 років для всіх Європейських міст з населенням більше ніж 50000 людей. Таким чином, міський атлас може використовуватися у якості засобу аналізу міста і порівняння міст в різних частинах Європи. Ці векторні шари надають дані для аналізу транспортної інфраструктури, навколишнього середовища та землекористування. Легенда атласу охоплює міське землекористування і бере за основу класифікацію CORINE [20]. Більш високе розрізнення у поєднанні з геопросторовою векторною інформацією дозволяє будувати велику кількість аналітичних продуктів, таких як щільність дерев у зонах міста, швидкість росту міста та статистику по забудовам з мінімальною одиницею картографування (Minimum Mapping Unit) 0.25 та 1 га.

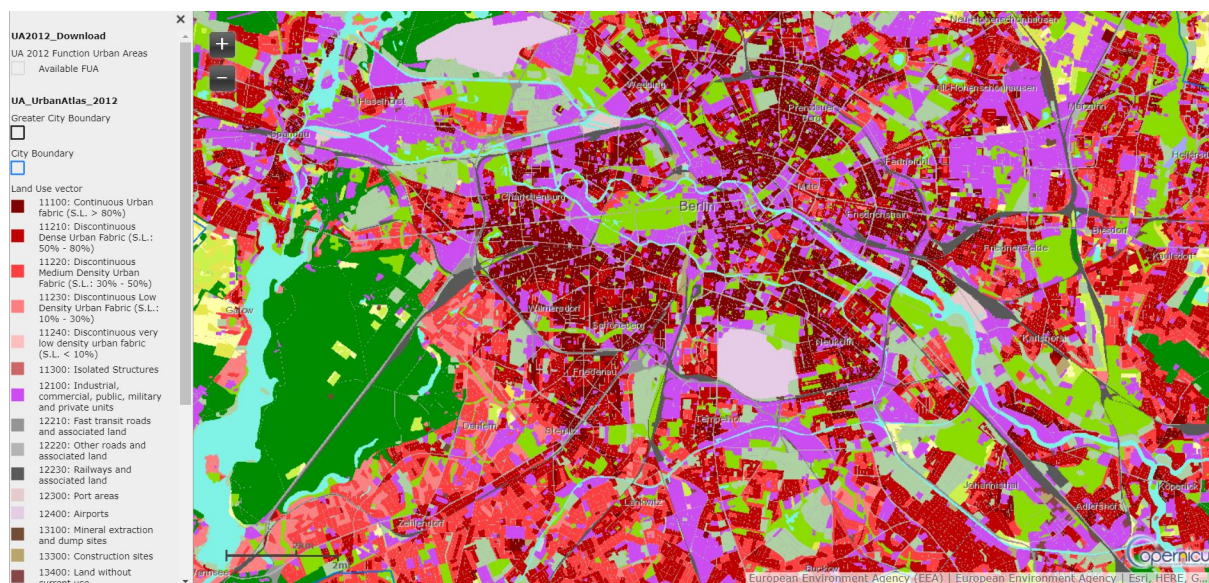


Рисунок 1.8 – Міський Атлас для міста Берлін за 2012-ий рік

Інститут Космічних Досліджень НАНУ та ДКАУ у проєкті Horizon-2020 ERA-Planet SMURBS розробляє сервіси моніторингу розумних міст [21] з використанням супутникових даних, в тому числі в ході цього проєкту був розроблений перший Український міський атлас [22] для міста Київ та

для міста Ірпінь (рис. 1.9).

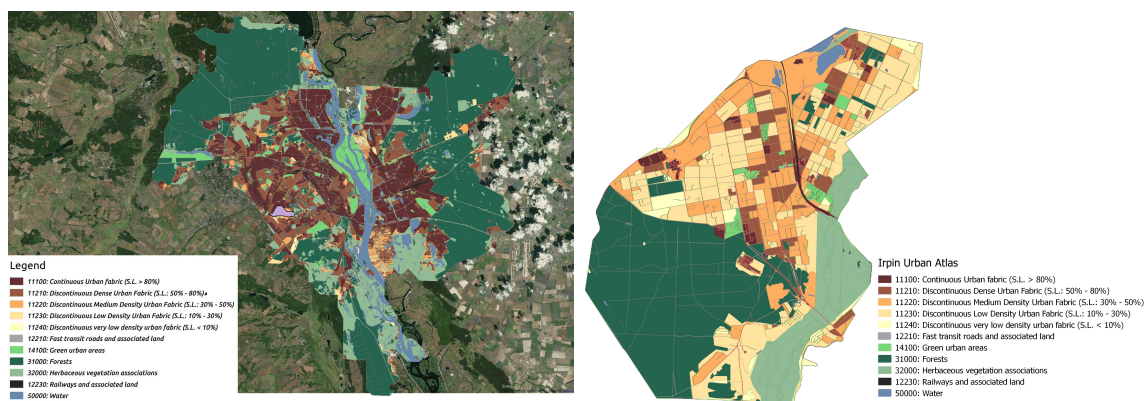


Рисунок 1.9 – Міський Атлас для міста Київ 2018 та Ірпінь 2019

Злиття растрових та векторних даних також можна вважати збільшенням просторового розрізнення, при якому супутникові характеристики поверхні землі в растровому вигляді усереднюються по об'єктам, що визначаються геопросторовою геометрією [25, 23, 24]. Також важливим аспектом у застосуванні такого методу є те, що просторове розрізнення в такому випадку буде вимірюватися не в одиницях метрів квадратних, а в одиницях площі мінімального об'єкта.

Висновки до розділу 1

Наявні супутникові місії NASA та ESA дають можливість проводити аналіз різних характеристик земної поверхні із застосуванням мультиспектральних даних отриманих з застосуванням різних сенсорів. Доступ до цих відкритих даних став дуже простим та зручним для розробників після появи таких сервісів як EOBrowser та Google Earth Engine. Проте задача поєднання супутникових мультиспектральних супутникових даних з різним просторовим та часовим розрізненням є досі актуальною, оскільки для розв'язання прикладних задач є потреба в застосування таких даних.

Геопросторовий аналіз супутникових даних може здійснюватися на

основі растрових та векторних продуктів. Проте, якщо говорити про специфіку моніторингу та аналізу розумних міст, для отримання інтуїтивно зрозумілого та зручного у використанні інтерактивного продукту необхідно застосовувати векторну сегментацію міста по типу міських атласів. Для Європейських країн ця інформація надається за допомогою сервісу Copernicus Land Monitoring Service Urban Atlas, а для України Інститут Космічних Досліджень НАНУ та ДКАУ розробив в рамках проєкту ERA-Planet SMURBS методологію побудови міських атласів на основі відкритих даних.

Поєднання різних геопросторових даних у векторній та растровій формі дають можливість отримати геопросторові продукти з вищим просторовим розрізненням ніж початкові дані і ефективно використовуватися для геопросторового аналізу.

2 МОДЕЛЮВАННЯ ТЕМПЕРАТУРИ ЗЕМНОЇ ПОВЕРХНІ НА ОСНОВІ СУПУТНИКОВИХ ДАНИХ

Сучасні супутникові місії дають великі об'єми даних на основі яких будуються інтелектуальні продукти, що застосовуються для реалізації прикладних досліджень. Температура земної поверхні, це важливий характеристика яка може бути використана для побудови сервісів моніторингу розумних міст, аналізу планування міст, а також виявлення пожеж. Всі ці застосування отримуються шляхом поєднання геопросторових даних та математичних моделей. У даному розділі описані основні 2 підходи які використовуються у моделюванні температури земної поверхні на основі супутникових даних.

2.1 Основні підходи моделювання темперартури земної поверхні на основі супутникових даних

Основою для моделювання температури земної поверхні є Thermal-Infrared канали супутникових знімків. Зазвичай сенсори, що використовуються для вимірювання інтенсивності електромагнітних хвиль у термальних діапазонах мають не менше двох каналів з різними діапазонами довжини хвилі. Це необхідно для використання точних алгоритмів, що враховують поглинання термальної енергії атмосферою. Для моделювання температури земної поверхні є два основних підходи на основі яких будуються всі можливі варіації алгоритмів: Radiative Transfer Equation [26] та Split Window Algorithm [27].

2.2 Split Window Algorithm

Split Window Algorithm є алгоритмом з найбільшою точністю моделювання температури земної поверхні та потребує для реалізації дані отримані із двох термальних супутникових каналів та додаткової інформації про стан атмосфери та характеристики земної поверхні. Алгоритм бере припущення, що випромінювання відбите від поверхні землі проходить через безхмарну атмосферу можна виміряти і яку надалі будемо називати Відбиттям на рівні вершини атмосфери (ТОА відбиття), можна виміряти за формулою:

$$B_i(T_i) = \varepsilon_i B_i(T_s) \tau_i + (1 - \varepsilon_i) \overleftarrow{R_{atm_i}} \tau_i + \overrightarrow{R_{atm_i}}$$

Де $B_i(T_i)$, це відбиття в каналі виміряне на вершині атмосфери в інфрачервоному спектрі T_i і відповідно $B_i(T_s)$. τ_i це коефіцієнт ефективності пропускання випромінювання через атмосферу для каналу i . B_i це функція Планка і $B_i(T_s)$ це відбиття від абсолютно чорного тіла із температурою T_s кельвінів. $\overleftarrow{R_{atm_i}}$ та $\overrightarrow{R_{atm_i}}$ це термальне випромінювання атмосфери, що направлене вгору та вниз. Ці коефіцієнти можна отримати тільки за допомогою вимірів в атмосфері та або за допомогою атмосферних моделей. Після отримання яскравісної температури T_i та T_j на основі ТОА з двох термальних каналів можна вирахувати температуру земної поверхні за формулою:

$$LST = b_0 + (b_1 + b_2 \frac{1 - \varepsilon}{\varepsilon} + b_3 \frac{\Delta \varepsilon}{\varepsilon^2}) \frac{T_i + T_j}{2} + (b_4 + b_5 \frac{1 - \varepsilon}{\varepsilon} + b_6 \frac{\Delta \varepsilon}{\varepsilon^2}) \frac{T_i - T_j}{2} + b_7 (T_i - T_j)^2$$

Де ε це коефіцієнт випромінювання поверхні усереднений для двох каналів, тобто $\varepsilon = \frac{1}{2}(\varepsilon_i + \varepsilon_j)$, а $\Delta \varepsilon$ це відповідно різниця коефіцієнтів випромінювання, тобто $\Delta \varepsilon = \varepsilon_i - \varepsilon_j$. b_k коефіцієнти для k від 0 до 7 це коефіцієнти моделювання які обраховуються на основі наземних вимірів температури земної поверхні. Детальніше про цей алгоритм та його

реалізацію для підрахунку температури земної поверхні на даних Landsat-8 описано у роботі Чен Ду [28]. Якщо говорити про операційне застосування, його реалізація є дуже складною оскільки потребує додаткових наземних вимірів, атмосферних вимірів та або додаткове моделювання коефіцієнтів.

2.3 Radiative Transfer Equasion

Radiative Transfer Equasion підхід використовує лише один термальний канал і є простим для реалізації сервісу моніторингу LST у наближеному до реального часі режимі. Підрахунок відбувається на основі коефіцієнту випромінювання та яскравісної температури одного термального каналу. Найпоширенішим та найбільш практичним в плані реалізації є алгоритм, що складається із 7-ми кроків [29], він схематично продемонстрований на рисунку 2.1.

Перший крок - підготовка супутникових знімків L2 рівня обробки за допомогою виконання атмосферної корекції, радіометричної корекції та маскування хмар. Після цього можна реалізовувати необхідні кроки підрахунку LST, такі як підрахунок NDVI та підрахунок яскравісної температури.

Другий крок - це підрахунок TOA за допомогою термального каналу супутникового знімку. TOA для 10-го каналу Landsat-8 обраховується за формулою:

$$TOA = M_L * B_{10} + A_L$$

Де M_L - 10-й діапазонний мультиплікативний коефіцієнт масштабування, що відповідає 10-ому каналу, а A_L - 10-й смуговий коефіцієнт додавання. Обидва коефіцієнта можна знайти у файлі метаданих MTL продукту Landsat-8 [30, 31]

Третій крок - це розрахунок яскравісної температури з використанням TOA, K_1 та K_2 це коефіцієнт конденсації теплової

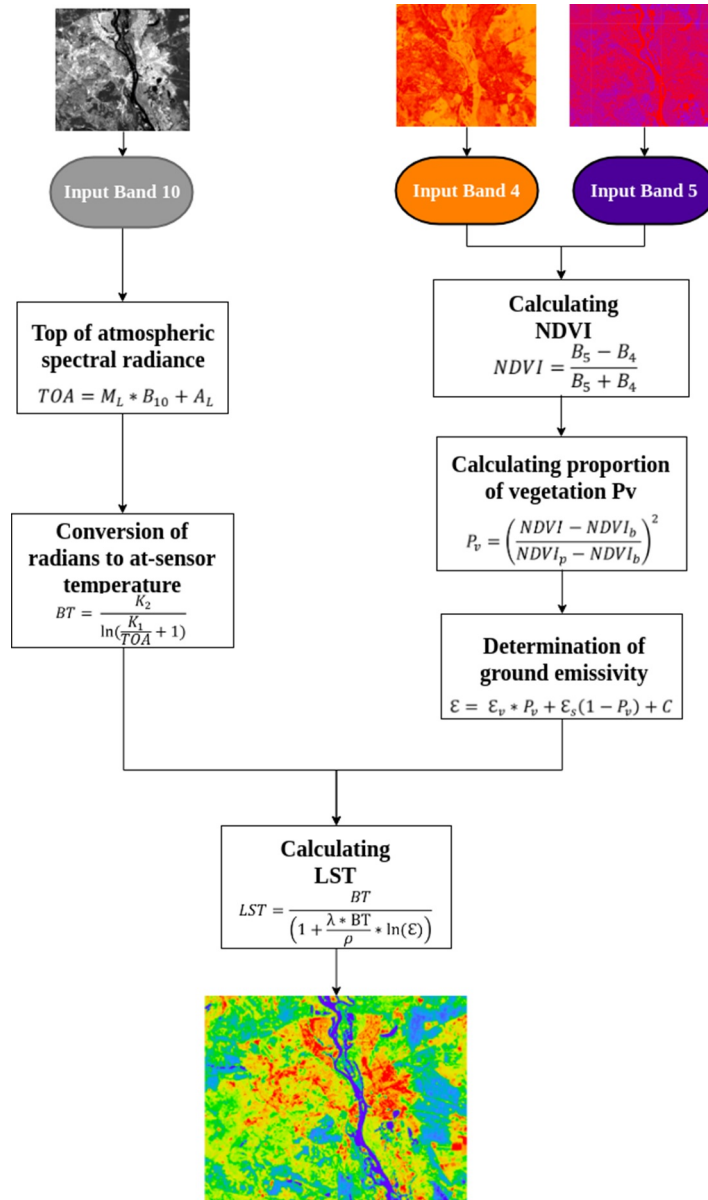


Рисунок 2.1 – Алгоритм підрахунку температури земної поверхні на основі Radiative Transfer Equation для знімків Landsat-8

конверсії для 10-го каналу знімку та коефіцієнт постійної теплової конверсії для 10-го каналу відповідно. Їх також можна знайти у файлі метаданих MTL продукту Landsat-8

$$BT = \frac{K_2}{\ln(\frac{K_1}{TOA} + 1)}$$

Після цього кроку ми отримуємо температуру, характерну для поверхні земної поверхні в градусах Кельвіна. Після цього нам необхідно

врахувати характеристики поверхні яка впливає на випромінювання та відбиття електромагнітних хвиль. Для цього можна використовувати 4-тий (B_4) та 5-тий (B_5) канали Landsat-8, тобто Red та Nir Red канали. Ми повинні обчислити індекс NDVI та пропорційний коефіцієнт рослинності P_v .

$$NDVI = \frac{B_5 - B_4}{B_5 + B_4}$$

Після цього пропорційний коефіцієнт вегетації може бути вирахованою з використанням значення $NDVI_b$ - значення NDVI голої землі та $NDVI_p$ - значення NDVI здорової рослинності. Вони можуть бути визначені як константи ($NDVI_b = 0.2$ та $NDVI_p = 0.5$ або можуть бути вираховані по обробленому знімку. У підході до обчислення цих характеристик по знімку $NDVI_p$ в може бути визначений як максимальне значення NDVI на знімку при відсутності карти земного покриву, при цьому точніше наближення до реального значення можна отримати лише з використанням карти земного покриву, підрахувавши середнє значення NDVI для класу покритої рослинністю землі. Складність отримання $NDVI_b$ полягає у врахуванні специфіки розташування зображення та зображених на ньому об'єктів, оскільки мінімальне значення NDVI не завжди відноситься до класу оголеної землі. Часто мінімальне значення індексу NDVI на зображенні належить водним об'єктам або хмарам, тому для правильного визначення цього значення необхідно мати деякі орієнтири: карта класифікації земного покриву або маска оголеної землі збудованої на основі Normalized Difference Bare Land Index. Ці дії дають можливість покращити результат підрахунку пропорційного коефіцієнту рослинності, але вони не є обов'язковими.

$$P_v = \left(\frac{NDVI - NDVI_b}{NDVI_p - NDVI_b} \right)^2$$

Для отримання коефіцієнту ефективності випромінювання теплової енергії земної поверхні, що передається через атмосферу, за допомогою розрахунку випромінюваної теплової енергії, використовуючи закон

Планка, ми повинні розрахувати емісію земної поверхні (ε)

$$\varepsilon = \varepsilon_v * P_v + \varepsilon_s(1 - P_v) + C$$

Де ε_v - емісія вегетації ($\varepsilon_v = 0.973$), ε_s - це емісія порожньої земної поверхні ($\varepsilon_s = 0.966$) і C - шорсткість поверхні (для плоскої поверхні $C=0$). Для деяких типів земельної поверхні можна вважати, що ε це константи. Ці константи є відомими і корисними для оптимізації обчислень. Їх отримання здійснюється за допомогою мультиспектральних продуктів супутнику MODIS [32]. Для водних об'єктів $\varepsilon = 0.991$, зазвичай NDVI водних об'єктів менше нуля. Якщо поверхня є голою землею $\varepsilon = 0.966$, зазвичай NDVI голої землі від 0 до 0.2. Якщо поверхня є штучною, тобто будівлі або дороги, тоді $\varepsilon = 0.962$, при цьому NDVI штучної поверхні має близькі значення до NDVI голої землі. Емісія поверхні з рослинністю розраховується за повною формулою, і найчастіше поверхня з рослинністю має $\text{NDVI} > 0.2$.

Як ми бачимо, для якісного розрахунку випромінювання необхідно використовувати карти класифікації земельного покриття. Якщо карти класифікації для вибраної території немає, можна зробити наступні наближення: випромінювання води, гола земля та будівлі (штучні об'єкти) - константи, випромінювання голої землі та штучних об'єктів рівні, тип поверхні можна визначити за індексом NDVI та визначити шорсткість поверхні як 0.

Останній крок - це розрахунок температури поверхні землі за допомогою температурної яскравості випромінювання поверхні землі у спектрі з довжиною хвилі λ , що для 10-го каналу Landsat-8 $\lambda = 10,892$.

$$LST = \frac{BT}{1 + \frac{\lambda * BT}{\rho} * \ln(\varepsilon)}$$

$$\rho = h * \frac{c}{\sigma} = 1.438 * 10^{-2} m$$

Де σ - це константа Больцмана, c - це швидість світла, h це стала Планка.

Для підрахунку температури земної поверхні на основі Landsat-8 використовується тільки 10-тий канал через те, що 11-тий канал був розроблений з дефектом, що спричиняє проблему направлених променів. Тобто промені світла, що направлені в сенсор по нормалі збільшує помилку фіксації інтенсивності випромінювання для 11-го каналу.

Висновки до розділу 2

Карти температури земної поверхні є важливою характеристикою земного покрову, яка використовується у багатьох прикладних задачах. Проте моделювання температури земної поверхні на основі супутникових даних не є тривіальною задачею. Існують два основних підходи для вирішення цієї задачі: Split Window Algorithm та Radiative Transfer Equation. Перший підхід використовує два термальних канали і дає найвищу точність моделювання, проте потребує додаткових наземних вимірів та атмосферну інформацію, що сильно ускладнює реалізацію цього підходу в операційних цілях. Другий підхід використовує тільки термальний канал ари цьому є дуже простим та зручним в реалізації, тому він був використаний як основний алгоритм, що використовується для подальших досліджень.

3 НЕЙРОННО МЕРЕЖЕВА МЕТОДОЛОГІЯ ЗБІЛЬШЕННЯ ПРОСТОРОВОГО РОЗРІЗНЕННЯ ТЕМПЕРАТУРНИХ ПРОДУКТІВ SENTINEL-3

Розроблено методологію збільшення просторового розрізнення супутникових карт температури земної поверхні Sentinel-3 на основі злиття мультиспектральних супутникових даних та застосування рекурентної нейронної мережі з архітектурою LSTM. Дана методологія була апробована на даних доступних для міста Київ і результат був порівняний із класичними регресійними методами, які використовуються для таких задач: лінійна регресія, баєсова регресія, випадкові ліси та багатошаровий перцептрон.

3.1 Рекурентні нейронні мережі

нейронні мережі є математичними моделями штучного інтелекту, які є реалізаціями примітивних моделей роботи мозку, яка заключається у наявності нейронів які отримують вхідні сигнали і на їх основі генерують вихідні сигнали. За допомогою штучних нейронних мереж розв'язується велика кількість прикладних задач у великій кількості наукових напрямків, починаючи від розпізнавання обличчя [33, 34], медичних цілях [35] до вирішення задач класифікації сільськогосподарських культур на основі супутникових даних [36, 37], прогнозування врожайності сільськогосподарських полів [38, 39] а також для обробки супутникових даних [40, 41] та постобробки супутникових продуктів [42].

Рекурентні нейронні мережі [43] це тип нейронних мереж, які мають зворотній зв'язок у нейронах, що дає можливість імітувати пам'ять нейрона. Таким чином вихід для кожного наступного входу нейрону залежить не тільки від входу нейрона, але також від минулого виходу. Ця

особливість дає нові можливості у розв'язуванні задач класифікації та регресії над даними які мають структуру часових рядів або мають часову залежність.

Стандартні алгоритми навчання рекурентної нейронної мережі [44] включає 7 кроків:

1) Подання вектору із навчального датасету, що має часову структуру

2) Нейронна мережа з ініційованими вагами та баєсами проводить вхід як сигнал здійснючи над ним обчислення за допомогою нелінійної складної функції. При створенні нової нейронної мережі можна використовувати випадкову ініціалізацію вагів.

3) Отримання виходу нейронної мережі у дискретній або недискретній формі в залежності від задачі, що розв'язується.

4) Обрахунок функції помилки. Часто функція помилки ґрунтується на ентропійних показниках розподілів вхідних та вихідних даних. Дані функції формально не є метриками відстані, оскільки вони не є симетричними.

5) Використовуючи метод зворотного розповсюдження помилки змінюються ваги в залежності від напрямку градієнту помилки.

6) Кроки 1-5 повторюються до завершення процесу навчання нейронної мережі. Зупинку можна здійснювати за допомогою правила рукава, методу Early Stopping або просто визначивши кінцеву кількість ітерацій навчання.

7) Використання навченої нейронної мережі для розв'язку задачі класифікації або регресії.

На рисунку 3.1 зображено роботу рекурентної нейронної мережі. Вхід X_i є вектором із m вимірів M_1, M_2, \dots, M_n .

Для кожного входу сигнал, що проходить через нейрон рекурентної нейронної мережі обчислюється вихід Y_i та інформація з входу H_i . З математичної точки зору вихід Y_i обраховується за такою формулою:

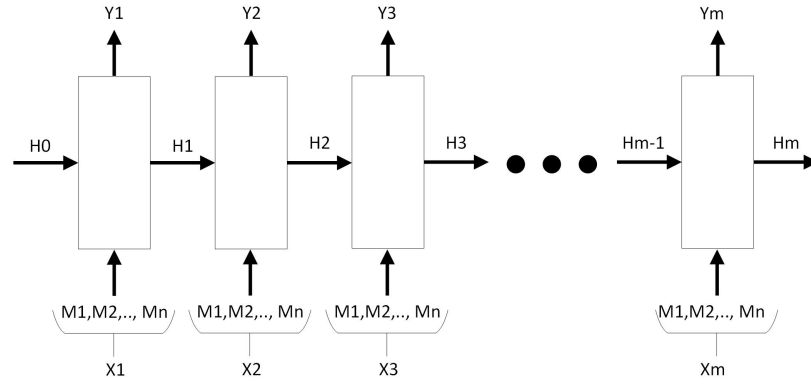


Рисунок 3.1 – Стандартна схема роботи рекурентної нейронної мережі

$$Y_i = f_{act}(W^{(hy)} H_i)$$

Де $W_{(hy)}$ це вагові коефіцієнти виходу, H_i це інформація входу, яка вираховується на основі інформації попереднього входу H_{i-1} , ваги рекурентного нейрону $W^{(hh)}$ та ваги вхідного нейрону $W^{(hx)}$:

$$H_i = \tanh(W^{(hh)} H_{i-1} + W^{(hx)} x_i)$$

f_{act} це нелінійна функція активації нейрону. Найпопулярнішими є функції:

$$Softmax(X) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

$$ReLU = \begin{cases} 0 & \text{при } x < 0 \\ x & \text{при } x \geq 0 \end{cases}$$

$$Sigmoid = \frac{1}{1 + e^{-x}}$$

В якості метрики помилки яка буде використовуватися для алгоритму зворотного розповсюдження помилок найчастіше використовують функцію cross-entropy яка вираховує різницю ентропії у розподілах випадкової множини на вході та на виході.

$$J^t = \sum i = 1n(y_{ti}' \log(y_{ti}))$$

3.2 Архітектура нейронної мережі з довгою та короткою часовою пам'яттю

Нейронна мережа з довгою та короткою часовою пам'яттю (LSTM) [45] це різновид рекурентних нейронних мереж, які здатні зберігати в нейронах інформацію про минулі входи на велику кількість ітерацій [46, 47]. Будь яка рекурентна нейронна мережа має форму ланцюжка повторюваних модулів рекурентних нейронів, що послідовно передають інформацію із свого входу на вхід наступного нейрона. Структура LSTM може представлятися так само у вигляді ланцюжка, проте внутрішньо кожен нейрон набагато більш складний з математичної точки зору, оскільки окрім функції активації додаткові функції для розповсюдження та зберігання інформації у вигляді стану нейрону. На рисунку 3.2 зображена схема роботи нейронної мережі LSTM.

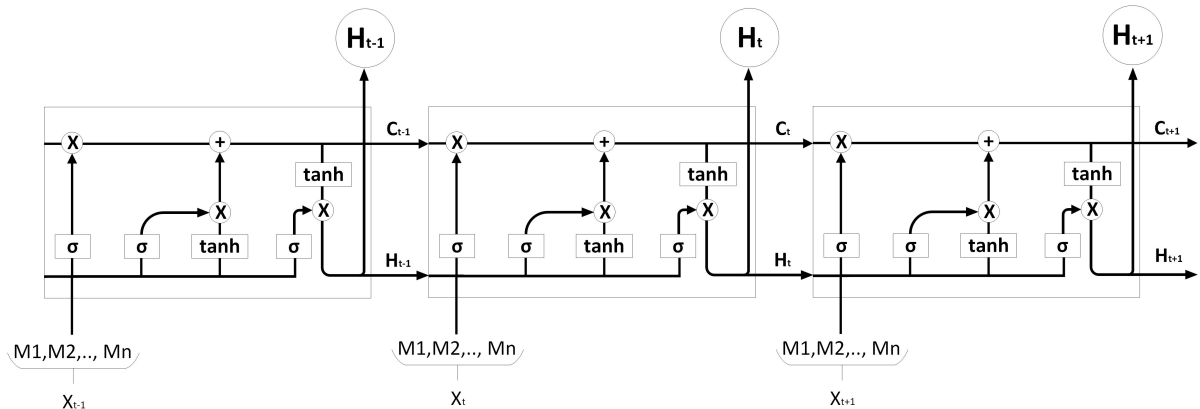


Рисунок 3.2 – Стандартна схема роботи LSTM

У даній схемі лінії показують напрямки проходження інформації у вигляді векторів. Ці вектори проходять через шари нейронної мережі, що навчаються, вони позначаються через прямокутники (σ , \tanh) та через векторні операції множення та додавання (X , $+$). Довгострокова пам'ять LSTM реалізується за допомогою стану нейронів (C_t , що передається від попереднього нейрону до наступного. Контроль інформації, що

зберігається в стані відбувається за допомогою так званих воріт (gates) що позначаються за як σ функції. Пропуск інформації відбувається за допомогою множення стану на вихід сігмоїдальної функції. Це можна вважати фільтруванням інформації по заданим умовам.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Сігмоїд повертає значення від 0 до 1 і це є коефіцієнтом частини інформації яку фільтр пропустить і збереже в стані мережі. 0 означає, що інформація через gate не пройде, 1 означає, що вся інформація пройде і буде збережена.

При отриманні нейроном входу та стану в першу чергу відбувається видалення інформації із стану, це відбувається за допомогою сигмоїдальної функції як і запам'ятовування. Цей фільтр називається фільтром забуття (forget gate layer). Цей фільтр для кожного числа із вектору стану повертає значення від 0 до 1 таким чином вибираючи скільки інформації необхідно забути. Функція фільтру забуття:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

Після цього вибирається нова інформація, що буде зберігатися у стані нейрону, це відбувається за допомогою шару вхідного фільтру (input layer gate), який є також сігмоїдальною функцією яка вибирає які значення вектору стану необхідно оновити і тангенціальна функція активації яка створює вектор нових значень стану. Добуток виходів сігмоїдної функції та тангенціальної функції після цього додається до стану мережі.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t$$

Далі відбувається обчислення виходу нейрону LSTM. Для цього на

основі минулого виходу та теперішнього входу відбувається обчислення сігмоїдальної функції після чого вихід цього шару множиться із виходом тангенціальної функції над станом нейрону та в результаті отримується вихід t -го нейрону LSTM h_t :

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

В сфері моніторингу землі зараз активно використовуються рекурентні нейронні мережі у задачах класифікації земного покриву, прогнозу врожайності та обробки супутникових продуктів. Архітектури рекурентних нейронних мереж стали дуже популярними у цій галузі через свою специфіку роботи з часовими рядами векторів, оскільки будь який набір супутникових даних можна інтерпретувати в такому вигляді. Проте для отримання коректного результату вхідні дані потребують гармонізації та формування часових рядів ознак інформативних для розв'язання поставлених прикладних задач.

3.3 Використання нейронної мережі Long Short Term Memory для збільшення просторового розрізнення температурних карт

Вхідними даними для моделі штучної нейронної мережі Long Short Term Memory є вектори, які є часовими рядами вимірів, тобто трьох вимірні вектори. Ці вектори формуються на основі пікселів NDVI та LST обрахованих на основі гармонізованих часових рядів супутникових знімків Sentinel-2, Landsat-8 та Sentinel-3 таким чином:

– Оптимальна кількість дат для роботи алгоритму дорівнює трьом. Перша дата це дата для якої відбувається збільшення просторового розрізнення даних і формується. Друга дата це минула дата

– Максимальна кількість дат для роботи алгоритму це 25 дат, при збільшенні кількості дат точність алгоритму починає падати. До того ж велика кількість дат призводить до виникнення прокляття великої розмірності даних, через що модель перенавчається і має в результаті дуже низьку дисперсію.

– Кожна дата містить в собі 4 виміри: NDVI даних високого розрізнення із набору даних Sentinel-2 та Landsat-8, LST даних високого розрізнення із набору даних Landsat-8 та збільшене розрізнення Sentinel-3, NDVI обрахований на даних Sentinel-3, LST обрахований на даних Sentinel-3.

– Для останньої дати в набору даних LST високого розрізнення замінюється на растр, який обраховується за формулою:

$$LST_{last} = \frac{\partial LST}{\partial t} * \Delta t + LST_{pre}$$

Де $\frac{\partial LST}{\partial t}$ це диференціал від функції залежності LST високого розрізнення від часу, Δt -це кількість днів між датою збільшення розрізнення та останім знімком LST з високим просторовим розрізненням LST_{pre} . Також це значення можна отримати за допомогою середньої на день різниці температур помноженої на кількість днів, доданої до останньої карти температури земної поверхні:

$$LST_{last} = \frac{LST_{t2} - LST_{t1}}{t2 - t1} * (t3 - t2) + LST_{t2}$$

Де $t3$ - це дата для якої відбувається моделювання температури земної поверхні, а $t1, t2$ це 2 послідовні минулі дати з наявними даними.

– Дані перед навчанням необхідно нормалізувати за допомогою мінімум максимум скейлу, тобто для вирахувати мінімальне та максимальне значення каналу для кожного виміру, створити таким чином вектор максимальних та мінімальних значень і після цього нормалізувати за формулою:

$$Data_{norm} = \frac{Data - V_{min}^-}{V_{max}^- - V_{min}^-}$$

Після ініціалізації нейронної мережі LSTM вона навчається робити прогноз на основі векторів у формі $N \times T \times 4$ повертати 1 число, яке інтерпритується як температура земної поверхні для визначеного дня. Тобто стек карт NDVI та LST за T дат є входом, а виходом є карта температури земної поверхні за дату T , для того щоб збудувати карту за дату $T+1$ необхідно видалити із стеку першу дату і додати дані за дату $T+1$. Для покращення роботи алгоритму навчену модель не потрібно ініціювати заново, достатньо її донавчити на нових даних, при цьому окремо виділяючи навчальні дані по пожежам та створюючи новий датасет який буде завжди додаватися до датасетів до кожного дня моделювання. Ця операція збільшить ймовірність розпізнати активний вогонь або підземну пожежу при збільшенні розрізнення супутникових знімків Sentinel-3.

Далі необхідно провести калібрування моделі на реальних прикладах з підбором гіперпараметрів. Експериментальним чином було отримано результат, що деякі параметри можна залишити статичними, а деякі необхідно підбирати в залежності від довжини входу. Статичними параметрами є функція втрат, на місце якої при ініціалізації можна вибрати середню квадратичну помилку, функцію активації визначити як гіперболічний тангенс, рекурентну функцію активації як сігмоїд, алгоритм оптимізації Adam, що використовує експоненційні середні значення ковзних вікон отримані на градієнті за поданим батчем та має швидкість навчання 0.01.

$$m_t = \beta_1 m_{t-1} + g_t - \beta_1 g_t$$

$$v_t = \beta_2 v_{t-1} + g_t - \beta_2 g_t$$

Де m_t та v_t це середні значення ковзних вікон для батчу під номером t , g_t це градієнт для батчу t , а β_1, β_2 це гіперпараметри алгоритму, які можна підбирати або використати стандартні значення $\beta_1 = 0.9$, а $\beta_2 = 1$ або 0.999. Для запобігання перенавчання можна встановити dropout параметр

з ймовірністю 25%, який реалізується як множення функції активації f_{act} на випадковий вектор D елементи якого розподілені за законом Бернуллі:

$$f_{act}^{new} = D * f_{act}$$

Таким чином ми можемо випадковим чином на кожному батчі навчальних даних вибиарати тільки частину нейронів, що буде навчатися. Такий спосіб регуляризації також сприяє підвищенню точності роботи алгоритму на різних входах, враховуючи, що повністю викінуті із навчальних даних пікселі пошкоджені хмарами та тінями є неможливим.

В якості архітектури LSTM було обрано схему "Всі в один". На рисунку 3.3 зображено дану нейронну мережу

Вибір гіперпараметрів для може здійснюватися за допомогою методів крос-валідації та Grid Search. Основним гіперпараметром який найбільше впливає на точність отриманого результату є кількість прихованих шарів нейронної мережі. Дані параметри можуть змінюватися для різних територій моделювання. Для невеликих територій або територій для яких 1 клас земного покрову є домінуючим, тобто кількість пікселів інших класів значно є відносно не суттєвою або відсутньою, можна використовувати одношарову архітектуру, або обмежити кількість шарів для перебору.

Даний метод є простим у масштабуванні і може бути використаний на рівні країни. Також важливою особливістю є те, що модель LSTM стійка до помилок у вхідних даних та дірок, тому відновлення хмар є не обов'язковим кроком, який можна пропустити, проте для отримання максимально високої точності збільшення розрізнення даних Sentinel-3 його варто виконувати. Також при збільшенні території можна збільшувати і частоту оновлення даних, оскільки на рівні країни наприклад Україна, для всієї території є неповне покриття даними Sentinel-2 та Landsat-8 щодня. Не повне покриття означає, що присутні дані тільки для однієї або декількох смуг. У такому випадку, можна

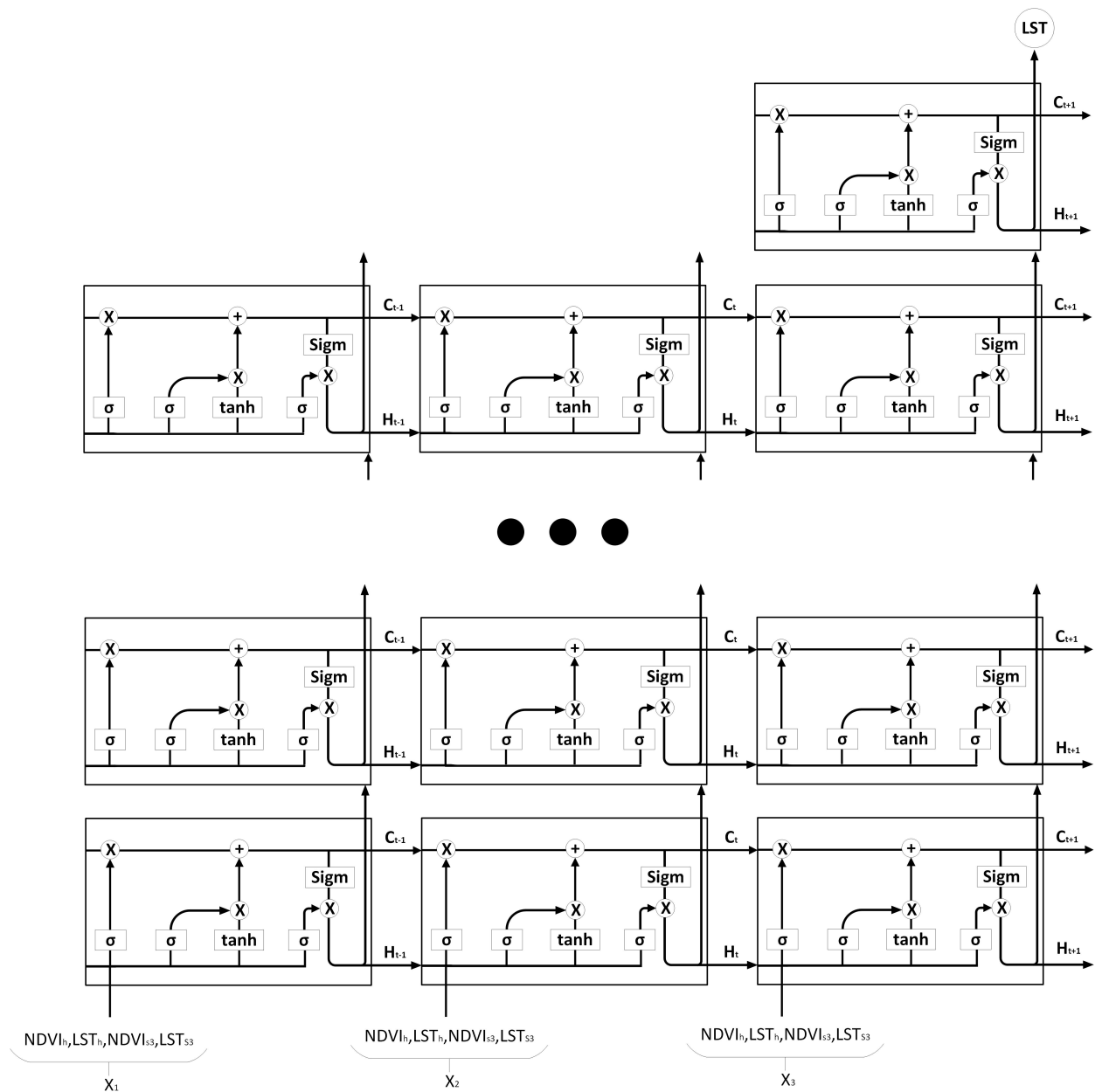


Рисунок 3.3 – Архітектура моделі LSTM для збільшення просторового розрізнення температурних карт

використати алгоритм відновлення хмар для моделювання покриття у зонах де відсутні знімки, таким чином покрити всю територію України. Процес відновлення пропущених полос очевидно не має такої високої точності як відновлення хмар, проте дає можливість будувати щоденну карту температури земної поверхні для території країни. Також для збільшення часового розрізнення можна використовувати інші супутникові дані більшого або рівного просторового розрізнення, проте

потрібно враховувати їхнє покриття, оскільки вони можуть мати дані не на кожній території. Саме тому вважається, що дані Sentinel-8 та Sentinel-2 є універсальним джерелом інформації, оскільки їхні місії гарантовано продовжуються і після запуску Landsat-9 та Sentinel-2C кількість покритих дат збільшиться. На рисунку 3.3 зображено структуру методу збільшення просторового розрізнення. температурних карт.

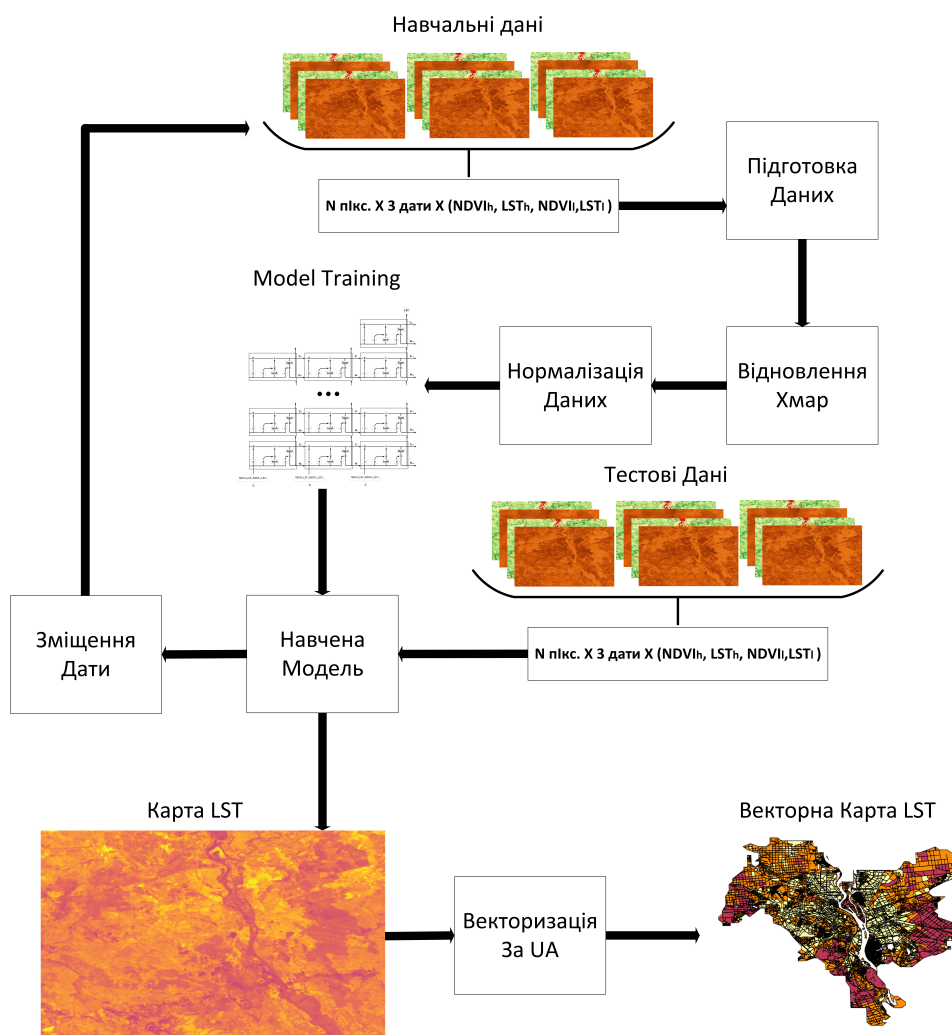


Рисунок 3.4 – Метод збільшення просторового розрізнення даних Sentinel-3

Даний метод може бути розбитий на 10 кроків:

1) Завантаження супутникових даних Sentinel-2, Landsat-8 та Sentinel-3. Та збереження їх у сховищі даних у зручному для подальшого використання стані.

2) Для оптичних даних Sentinel-2 та Landsat-8 необхідно провести атмосферну корекцію, радіометричну корекцію, сформувати маску хмар та замаскувати хмари, привести до 30-ти метрового просторового розрізнення та зберегти у типі Float32.

3) Для всіх даних необхідно провести відновлення хмар за методом описаним у розділі 4.2, або за допомогою будь-якого іншого доступного методу відновлення хмар.

4) Ці супутникові дані вчитуються у формі часових рядів із 4-ох вимірів: $NDVI_h$ - NDVI високого просторового розрізнення, LST_h - температура земної поверхні високого просторового розрізнення, $NDVI_l$ - NDVI низького просторового розрізнення отриманий із знімків Sentinel-3, LST_l - температура земної поверхні з низьким просторовим розрізненням.

5) Сформувати навчальний та тестовий набори даних. Якщо необхідно провести збільшення просторового розрізнення для дати t_4 , тоді навчальний датасет формується для трьох послідовних доступних дат t_1, t_2, t_3 і навчати на доступній карті температури земної поверхні t_3 , а тестування необхідно здійснювати для дат t_2, t_3, t_4 . При цьому для тестування необов'язково мати наявну карту температури земної поверхні.

6) Нормалізація отриманих даних за допомогою Мін-Макс нормалізації.

7) Ініціалізація та калібрування нейронної мережі або ініціалізація нейронної мережі за вже збереженими вагами відкаліброваної нейронної мережі.

8) Збільшення просторового розрізнення карти температури земної поверхні за тестовим набором, що використовується у роботі з нейронною мережею.

9) Зміщення дати для побудови навчального набору даних для наступної ітерації на основі нової карти температури земної поверхні та нових вхідних даних.

10) Векторизація карти земної поверхні для отримання інтерактивної

карти для подальшого аналізу.

Висновки до розділу 3

Рекурентні нейронні мережі, це особливий тип нейронних мереж, який використовує зворотній зв'язок у нейронах для симуляції пам'яті, таким чином додаючи зв'язок між входами для різних нейронів. Ця особливість створює перевагу даних типів нейронних мереж над іншими типами у задачах в яких використовуються дані, що мають часову структуру. Оскільки супутникові дані у більшості прикладних задач використовуються як часові ряди мультиспектральних даних, дослідження використання даного типу нейронних мереж має високу перспективу. Найбільш популярною архітектурою рекурентних нейронних мереж є мережа з довготривалою та коротко тривалою пам'яттю (LSTM), яка надалі стане основою для методу збільшення просторового розрізнення супутникових даних Sentinel-3.

Було розроблено методологію збільшення просторового розрізнення супутникових карт температури земної поверхні Sentinel-3 на основі математичної моделі нейронної мережі LSTM з архітектурою "Всі в Один". Дана методологія бере на вхід часові ряди карт NDVI та LST за минулі дати та поточну дату. Особливістю методології є формування входу для дня моделювання з використанням обрахунку синтетичної карти температури земної поверхні на основі диференціалу функції залежності температури земної поверхні від дня, отриманої при аналізі часових рядів температурних карт. Таким чином дану методологію можна використати для збільшення просторового розрізнення карт температури земної поверхні для дат при яких немає даних Landsat-8, на основі даних Sentinel-2. Таким чином з'явилася можливість отримувати температурні карти земної поверхні кожні 2-3 дні з 30-ти метровим просторовим розрізненням. Тобто у двічі вдалося збільшити часове розрізнення покриття картами температури земної поверхні для міста Київ.

4 ВИКОРИСТАННЯ МЕТОДОЛОГІЇ ЗБІЛЬШЕННЯ ПРОСТОРОВОГО РОЗРІЗНЕННЯ СУПУТНИКОВИХ КАРТ ТЕМПЕРАТУРИ ЗЕМНОЇ ПОВЕРХНІ SENTINEL-3

В даному розділі описано процес завантаження та обробки супутникових даних Sentinel-2, Landsat-8 та Sentinel-3. Які використовуються для створення нової колекції супутникових даних температури земної поверхні на основі Sentinel-3 з високим просторовим та часовим розрізненням.

Розроблено алгоритм відновлення хмар на супутникових продуктах вегетаційних індексів NDVI та LST, який має суттєву перевагу через вбудований пошук аномальних показників у пікселях карт, які прибирають дефекти викликані не ідеальною маскою хмар та тіней.

Сформовано навчальний та тестовий датасет для збільшення просторового розрізнення температурних карт Sentinel-3, а також проаналізовано розподіл даних за допомогою емпіричного статистичного правила 68-95-99.7, тесту Колмогорова-Смірнова та тесту Шапіро-Уїлка.

Розроблена модель збільшення просторового розрізнення супутникових даних температури земної поверхні протестована на зібраних та проаналізованих даних, порівняно із класичними регресійними методами, які використовуються для таких задач: лінійна регресія, баєсова регресія, випадкові ліси та багатошаровий перцептрон та отримано за допомогою неї результат для міста Києва.

4.1 Підготовка та гармонізація супутникових даних

Дані супутникової місії Sentinel-2 доступні для завантаження на сайті Європейського космічного агентства порталі Copernicus Open Access Hub [48] і має основні 2 типи рівні обробки, це L1 рівень, тобто продукти

без обробки та L2 рівень, тобто оброблені продукти. Якщо говорити про операційне використання даних Sentinel-2, тоді завантажуються в першу чергу продукти рівня L1, які надалі проходять обробку локально, оскільки продукти L2 не є доступними в день зйомки. Основними фазами обробки продукту L1 є: атмосферна корекція, радіометрична корекція та цирус корекція. Для виконання даних операцій обробки можна використовувати відкриті бібліотеки та програмне забезпечення яке розробило Європейське Космічне Агенство, це відкрите програмне забезпечення має назву Sen2Core [49]. Після проведення даних стадій обробки генерується SLC файл з сегментацією основних класів зображених об'єктів на знімку, в який також входить маска хмар та тіней. Необхідно замаскувати всі хмари і тіні на знімку. Після цього необхідно з'єднати в один знімок канали червоного, синього, зеленого та близького інфрачервоного спектрів і перевести знімок в проекцію EPSG:32638 з просторовим розрізненням 30 метрів. Для цього можна використати білінійний метод перекомпонування.

Дані супутникової місії Landsat-8 надаються NASA, і їх так само необхідно обробити. Для обробки необхідно використати алгоритм атмосферної корекції FMask [50], розроблений під особливості для сенсорів місії Landsat-8. В результаті роботи алгоритму також генерується канал якості який використовується як маска хмар та тіней і за допомогою якої необхідно провести маскування цих дефектів. Після отримання даних Sentinel-2 та Landsat-8 знімки необхідно гармонізувати [51], скласти в один часовий ряд та підрахувати вегетаційний індекс NDVI.

Супутникові дані Landsat-8 та Sentinel-2 добре доповнюють один одного. Хоча дата повного повторення для першого 16 днів, для другого 5 днів, враховуючи перетин смуг Landsat-8 для деяких територій, таких як Київ, час повторної зйомки може дорівнювати 7-8 днів. Таким чином на рисунку 4.1 зображено покриття міста Києва даними супутникових знімків Landsat-8 та Sentinel-2 з 01.06.2019 по 01.09.2019, тобто за все літо.

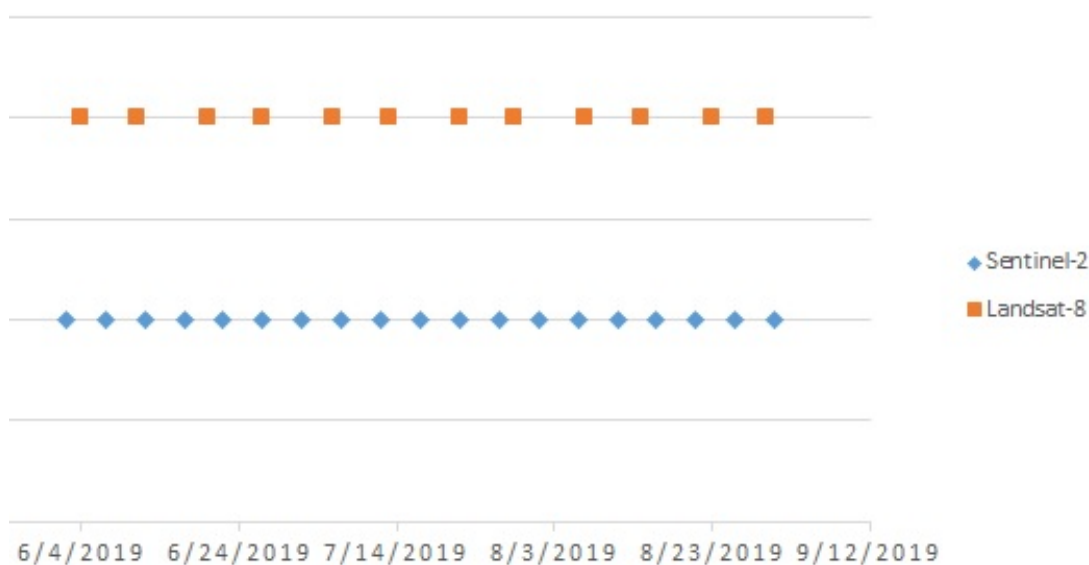


Рисунок 4.1 – Часове покриття міста Києва даними Sentinel-2 та Landsat-8 за літо

Таким чином на кожні 10 днів доступно 3-4 знімки гармонізованого Sentinel-2 та Landsat-8. На рисунку 4.2 зображено часовий розподіл знімків з 11.06.2019 по 20.06.2019. Де видно, що об'єднавши датасет Sentinel-2 та Landsat-8 кількість доступних знімків збільшується у двічі.

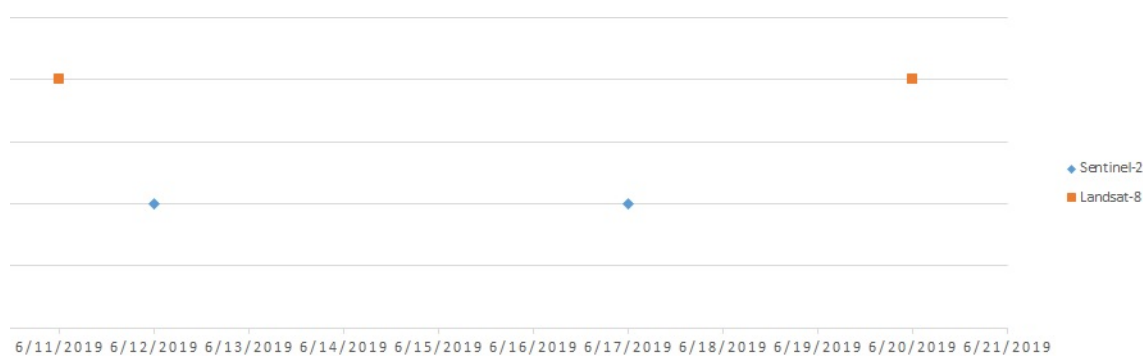


Рисунок 4.2 – Часове покриття міста Києва даними Sentinel-2 та Landsat-8 за 10 днів

Дані Sentinel-3 доступні так само як і дані Sentinel-2 на Copernicus Open Access Hub. Тому завантаження даних для Sentinel реалізується

просто через відкриту бібліотеку `sentinelhub-py` [52] на мові програмування Python. Ці дані мають специфічну структуру і зберігаються у вигляді архіву з NetCDF файлами та XML файлом в якому зберігається структура даних, що називається `xfdumanifest`. Відкриваються ці дані через відкрите програмне забезпечення Європейського Космічного агентства, яке називається SNAP [53] з підключенням бібліотек для роботи з Sentinel-3. 300 метрові продукти з каналами видимого і невидимого спектрів знімається за допомогою Ocean Land Color Instrument (OLCI) та зберігаються в EFR продуктах (Повне розрізнення TOA відбиття). 1 км продукти температури земної поверхні вимірюються за допомогою Sea Land Surface Temperature Radiometer (SLSTR) і зберігаються у LST продуктах другого рівня обробки. Граф для обробки даних Sentinel-3 є досить простим, він наведений на малюнку 4.3.

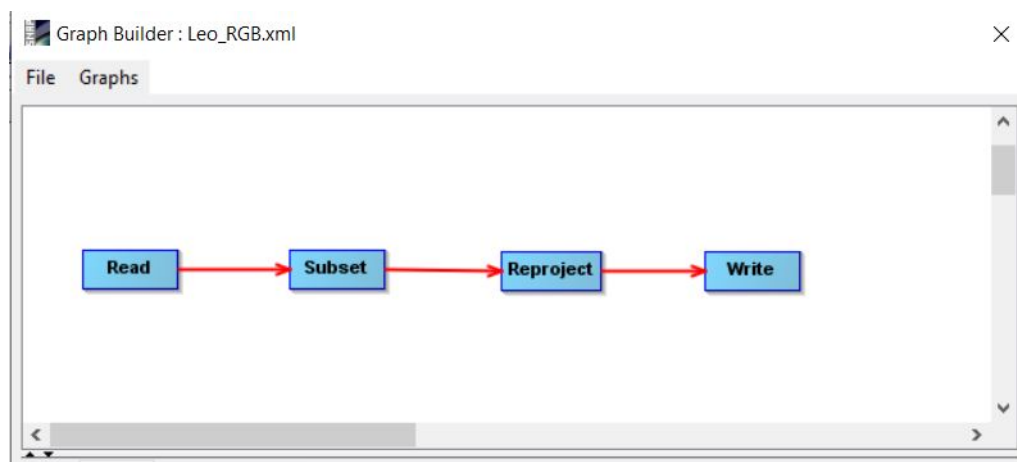


Рисунок 4.3 – Граф для обробки даних Sentinel-3

Формування таких графів для продуктів Sentinel дозволяють автоматизувати процес обробки та завантаження даних Sentinel-3, оскільки використовуючи SNAP Python API можна використовувати дані графи при побудові скриптів. Після зчитування даних необхідно зробити Subset, тобто вибрати підгрупу каналів для подальшої роботи. Для OLCI необхідно вибрати канали Oa4, Oa6, Oa8 та Oa17 які інтерпретуються як

червоний, зелений, синій та близький інфрачервоний, а для температурних продуктів SLSTR відповідно необхідно брати LST канал, при цьому перед Subset необхідно провести операцію Band Calculation та замаскувати всі пікселі які на каналі якості відмічені як хмари або тіні. Після цього необхідно перевести дані у необхідну для роботи проекцію за допомогою білінійного методу перекомпонування та зберегти у вихідному файлі GEOTiff формату. Після цього на основі OLCI знімків необхідно підрахувати індекс вегетації NDVI та перекомпонувати NDVI та LST карти Sentinel-3 у 30-ти метрове просторове розрізнення, для того щоб їх було просто поєднувати із даними Landsat-8 та Sentinel-2.

4.2 Відновлення хмар на супутникових продуктах

Захмареність знімків є головною проблемою використання оптичних даних. Для того щоб використовувати оптичні супутникові дані для проведення аналізу великих територій або для проведення моніторингу стану поверхні необхідно проводити відновлення хмар на супутникових даних. На малюнку 4.4 наведено приклад карти температури земної поверхні отриманої на основі знімків Landsat-8 за 07.06 2019 .

Задачу відновлення хмар на супутникових даних можна звести до задачі регресії. При цьому варто зазначити, що окрім дірок в даних хмари спричиняють також дефекти на знімку пов'язані із недосконалістю маски хмар та тінями, які є складними об'єктами для детектування. Тому оптимальний спосіб отримання безхмарних даних високої якості є моделювання повного знімку використовуючи як референсну інформацію непошкоджені пікселі знімку для якого відбувається відновлення хмар. Оптимальним алгоритмом для цього є алгоритм випадкових лісів, який працює у режимі регресії. Для цього використовуються дві функції для

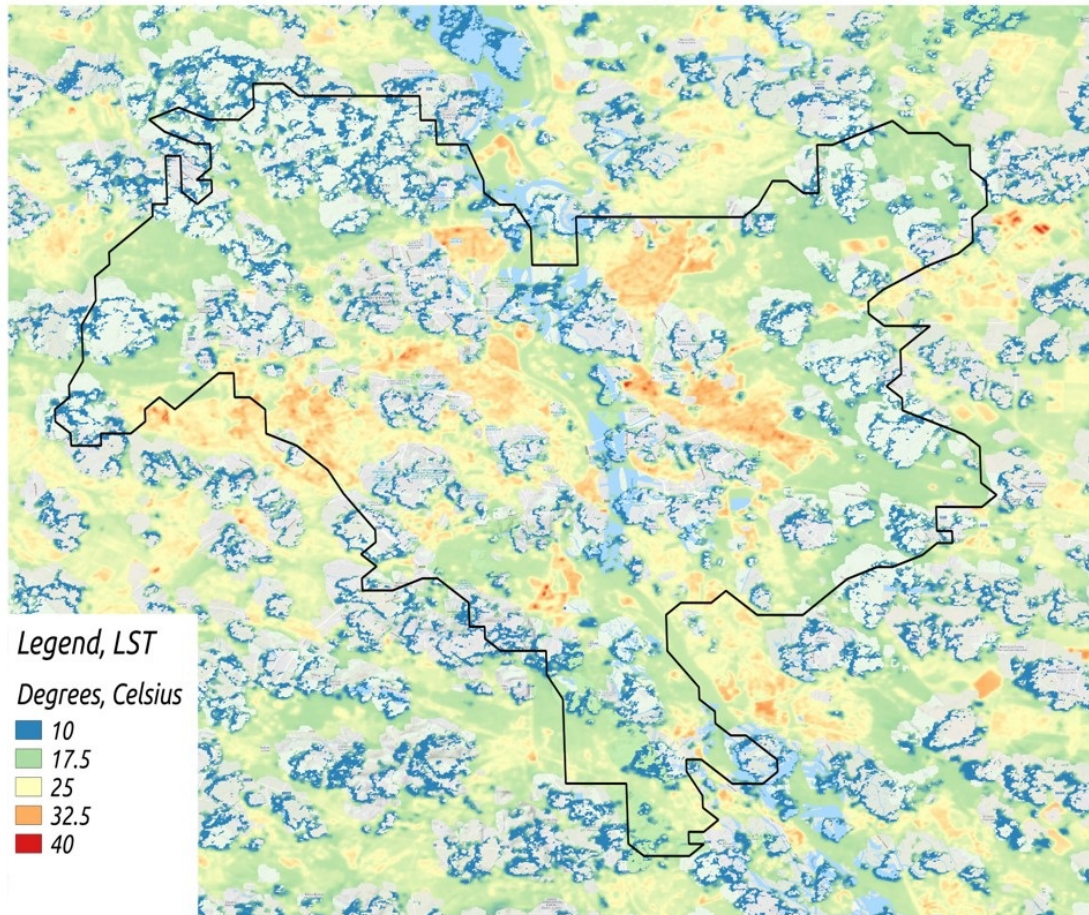


Рисунок 4.4 – Карта температури земної поверхні міста Києва після маскування хмар на основі даних Landsat-8 за 07.06.2019

розгалуження при побудові лісу на не дискретних значеннях:

$$MSE = \frac{1}{N} \sum_{i=0}^N (y_i - \mu)^2$$

$$MAE = \frac{1}{N} \sum_{i=0}^N |y_i - \mu|$$

У даній задачі обидві функції дають тотожний результат, тому в подальшому використовується функція MAE. Далі алгоритм на основі методу Джині обраховує важливість розгалужень за формулою:

$$ni_j = w_j C_j - w_{left(j)} C_{left(j)} - w_{right(j)} C_{right(j)}$$

Де ni_j - важливість розгалуження j , w_j - це взважена кількість семплів які

досягаються при розгалуженні j , C_j - значення домішків j , $left(j)$ - це ліве розгалуження для розгалуження j , $right(j)$ - це праве розгалуження для розгалуження j . І важливість для кожної ознаки після цього може бути вирахована за формулою:

$$fi_i = \frac{\sum_j ni_j}{\sum_k ni_k}$$

Де fi_i - це важливість ознаки i , ni_j - важливість розгалуження j , J - це всі розгалуження, що здійснюються на основі ознаки i .

При цьому часто для автоматичного розгалуження використовується так званий приріст інформативності ($Gain(T, X)$), який вираховується для T -цільові змінні та X - знака для розгалуження на основі функції ентропії у випадку класифікації, або MSE/EMU у випадку регресії:

$$Entropy = \sum_i i = 1^C - f_i \log(f_i)$$

$$Gain(T, X) = Entropy(T) - Entropy(T, X)$$

Відновлення хмар відбувається спочатку для продукту NDVI. Варіація цього індексу не сильно змінюється в невеликому проміжку часу, а для деяких типів земного покрову, таких як вода, штучна поверхня та гола земля може вважатися сталою. Ідея використання алгоритму заключається в тому, що дерева рішень перенавчаються на вхідних даних та моделюють вхідний знімок навчаючись на ньому ж. Після першої ітерації алгоритм видає карту NDVI без дірок, проте оскільки на вхід при навчанні потрапили також пошкодженні хмарами та тінями дані, які не потрапили у маску хмар та тіней, перший вихід використовується для виявлення цих пошкоджених даних. Критерієм відкидання даних є:

$$|Y - Y'| > median_{T(Y)}(|Y_{T(Y)} - Y_{T(Y)}'|)$$

Де Y та Y' це значення пікселів на знімку до та після відновлення. Модуль різниці значення пікселя до та після відновлення порівнюється з

медіаною модуля різниці значення пікселів до та після відновлення для всіх пікселів які відносяться до однакового з Y типом земного покрову. Тобто функція $T(Y)$ це функція класифікатор значення Y по типах земного покрову. Вона може бути реалізована як окрема функція яка в залежності від значення Y відносить його до класу земного покрову, або може бути замінена картою класифікації земного покрову. Враховуючи, що тип земного покрову змінюється дуже рідко і протягом року може вважатися константою, тоді можна спростити цей процес за рахунок використання глобальних [54, 55] або локальних карт земного покрову [56]. Після цього навчається регресія на оновленому датасеті та отримується фінальний датасет. Вхідними даними є доступні за місяць карти NDVI з проведеним Uniform filtering на них. Таким чином отримується безхмарна карта NDVI. На рисунку 4.5 зображено результат відновлення хмар для знімку Landsat-8 за 07.06.2019.

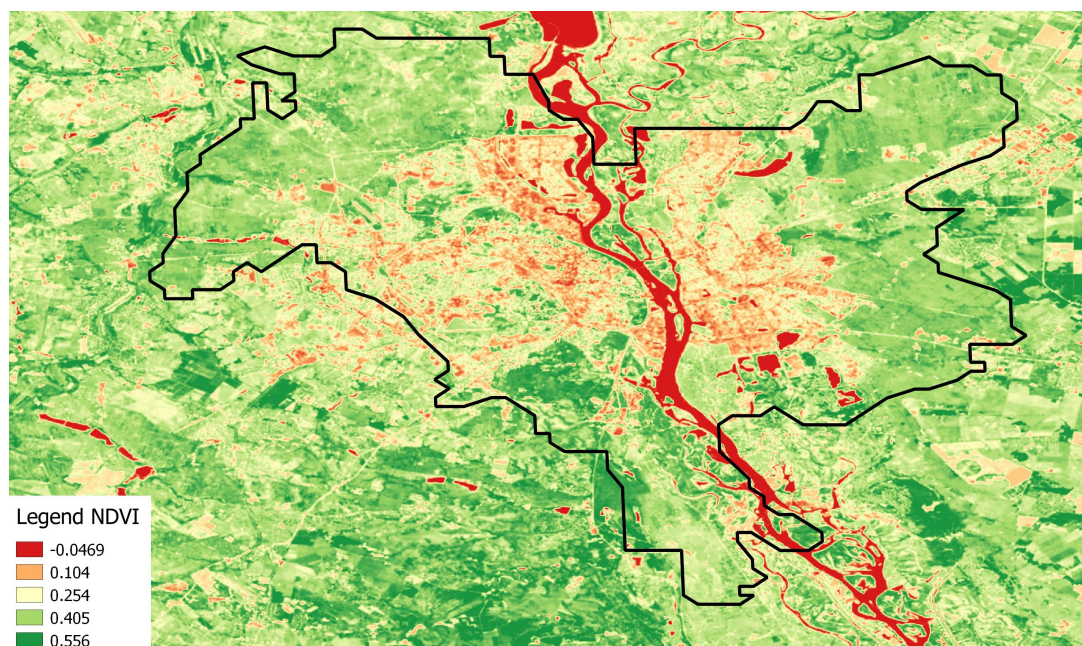


Рисунок 4.5 – Відновлена карта NDVI на основі знімків Landsat-8 за 07.06.2019

У таблиці 3.1 наведено MSE для різних діапазонів NDVI. Підрахунок помилки відбувається на реальних даних, тому отримання

нуля є неможливим, при цьому отримана точність задовольняє вимоги до вхідних даних.

Таблиця 4.1 – Середня помилка відновлення NDVI за діапазонами значення NDVI

Діапазон NDVI	MSE
$-1 < \text{NDVI} < 0$	0.0004
$0 \leq \text{NDVI} < 0.15$	0.0005
$0.15 \leq \text{NDVI} < 0.3$	0.0012
$0.3 \leq \text{NDVI} < 0.5$	0.0008
$0.5 \leq \text{NDVI} < 1$	0.0023

Після відновлення карти NDVI на основі неї відбувається моделювання температури земної поверхні, проте оскільки хмари пошкодили дані отримані із термального каналу, а саме термальну яскравість, застосовується такий же підхід для відновлення термальних карт. Для цього використовується цей же алгоритм, проте до входів окрім NDVI за місяць додається також LST за місяць. Таким чином отримується нова карта температури земної поверхні. На рисунку 4.6 наведено відновлену карту температури земної поверхні за 07.06 2019.

У таблиці а3.2 наведено MSE для відновлених даних. Найвища помилка при відновленні виникає при високих температурах земної поверхні, це пояснюється тим, що ці температури досягаються рідко і знаходяться у окремих зонах, що вважаються місцями температурних аномалій. Ці аномалії викликані особливостями поверхні землі в тому числі штучних поверхонь, які абсорбують та утримують багато тепла.

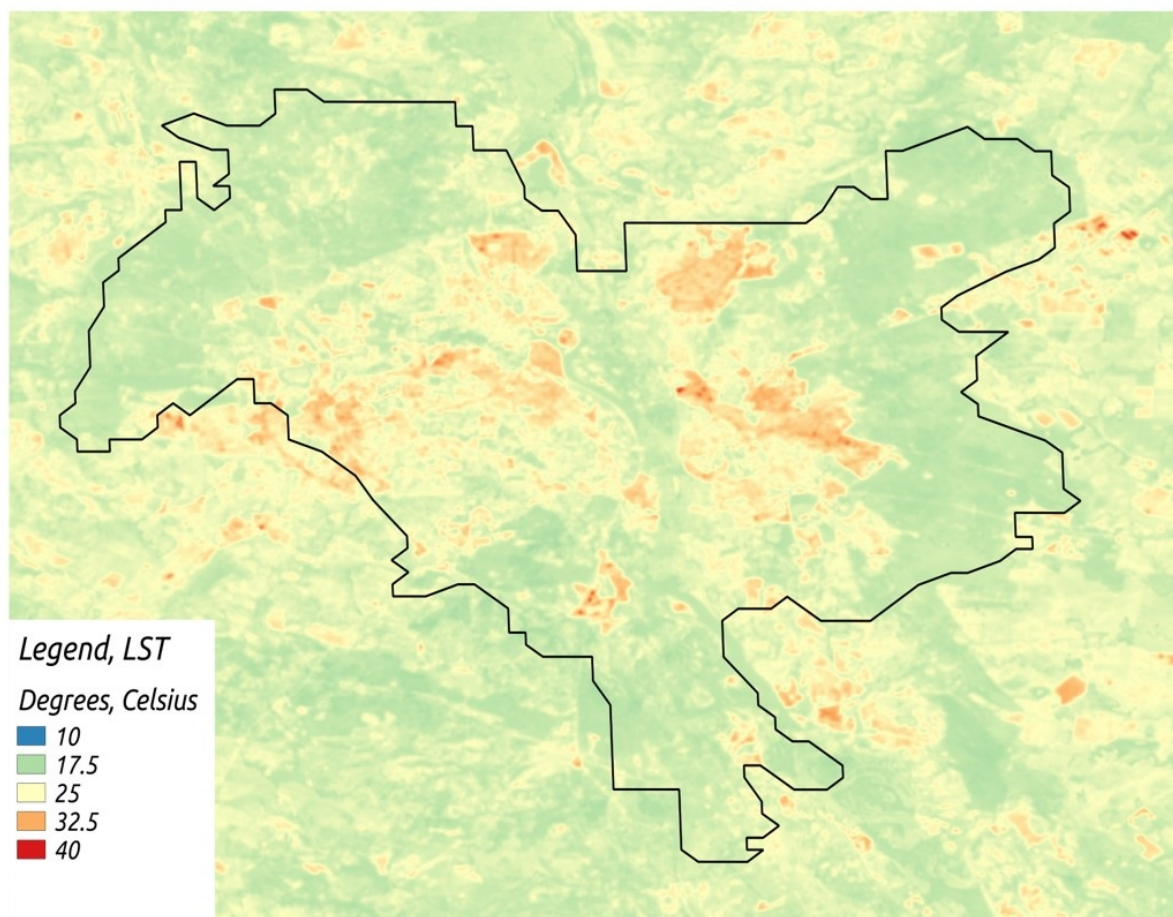


Рисунок 4.6 – Відновлена карта температури земної поверхні на основі знімків Landsat-8 за 07.06.2019

Таблиця 4.2 – Середня помилка відновлення температури земної поверхні за діапазонами температури земної поверхні

Діапазон LST	MSE
$20 < \text{LST} < 27$	1.8
$27 \leq \text{LST} < 31$	0.78
$31 \leq \text{LST} < 36$	1.4
$36 \leq \text{LST} < 40$	3
$40 \leq \text{LST} < 50$	4

4.3 Формування навчальних і тестових даних для моделі збільшення просторового розрізнення супутникових даних

Метод збільшення просторового розрізнення супутникових карт температури земної поверхні ґрунтується на моделі машинного навчання і тому необхідно провести аналіз вхідних даних, для того щоб зрозуміти який розподіл даних, чи підходить він для навчання та валідації моделі. При цьому потрібно зауважити, що основним режимом роботи моделі є отримання часових рядів супутникових даних з високим часовим розрізненням, тобто на вхід у реальному застосуванні будуть подаватися змодельовані дані. А валідація заключається на тому, що кожні 16 днів ми маємо точні карти температури земної поверхні отримані із супутника Landsat-8. Передбачається також донавчання моделі з кожною новою датою. Разом з цим донавчання моделі на нових датах та використання дат з низькою часовою різницею забезпечує високу точність яка буде коливатися випадковим чином, в залежності від захмареності даних та точностей відновлення даних у відповідні дні, між мінімальною точністю та максимальною точністю яку неможливо визначити одночасно. При цьому мінімальну точність можна оцінити при першому запуску моделі. При цьому запуску використовуються дані Landsat-8 в якості даних високого просторового розрізнення, оскільки дані Sentinel-2 не містять карти температури земної поверхні високого просторового розрізнення. Для першого навчання використаємо такі дані:

1) Першою датою є 3 травня 2019-го року. За цей день були наявними знімки RGBN та LST високого просторового розрізнення з 13 % хмарністю і знімки RGBN та LST Sentinel-3 із 30% захмареністю.

2) Дрегою датою є 19 травня 2019-го року. За цей день були наявними знімки RGBN та LST високого просторового розрізнення з 5% хмарністю і знімки RGBN та LST Sentinel-3 із 26% хмарністю.

3) Поточною датою моделювання є 4 червня 2019-го року для якої

були наявними знімки RGBN та LST з 42% хмарністю високого просторового розрізнення і знімки RGBN та LST Sentinel-3 з 45% хмарністю. При цьому знімок LST використовується як справжнє значення при навчанні, а в якості входу відбувається обрахунок штучного знімку LST на основі функції температури земної поверхні.

Для тестування використовується моделювання дати 20 червня 2019-го року для якої були наявні знімки високого просторового розрізнення RGBN та LST з 25% хмарністю та RGBN LST знімки Sentinel-3 з 23% хмарністю. Дані LST високого просторового розрізнення використовується як тестовий набір даних, а для поточного дня моделювання LST високого просторового розрізнення вираховується на основі минулих днів.

На рисунку 4.7 зображено приклад вхідних даних для одного дня 04.06.2019, який складається із карти NDVI високого просторового розрізнення (NDVI h), LST високого просторового розрізнення (LST hr), NDVI Sentinel-3 (NDVI lr) та LST Sentinel-3 яке збільшує своє просторове розрізнення до високого просторового розрізнення (LST lr). Для цих ознак проведено кореляційний аналіз за допомогою коефіцієнту Пірсона:

$$r_{xy} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2 \sum_{i=1}^m (y_i - \bar{y})^2}} = \frac{cov(x, y)}{\sqrt{s_x^2 s_y^2}}$$

Кореляційний аналіз показав високу кореляцію між картами температури земної поверхні високого просторового розрізнення та низького просторового розрізнення. Для 04.06.2019 кореляція температур дорівнює 0.77. При цьому кореляція між NDVI високого і низького просторового розрізнення є середньою і для цього дня дорівнює 0.56. Це пояснюється тим, що температура земної поверхні навідміну від NDVI має плавнішу зміну у переходах від одного пікселя від іншого і має краще усереднення значення по пікселю. Якщо говорити про кореляцію між NDVI та температурою вона рівна 0.55 та 0.40 для високого та низького просторового розрізнення. Тобто інформації NDVI не достатньо для відновлення температури земної поверхні. Якщо орієнтуватися на

алгоритми моделювання температури земної поверхні тоді для повного відновлення необхідно мати NDVI та яскравісну температуру і найкраще температура земної поверхні низького просторового розрізнення разом з NDVI низького просторового розрізнення інтерпретуються як дані, що містять інформацію про яскравісну температуру, а NDVI високого просторового розрізнення допомагає перерозподілити яскравість пікселів високого просторового розрізнення через коефіцієнт емісії пікселя.

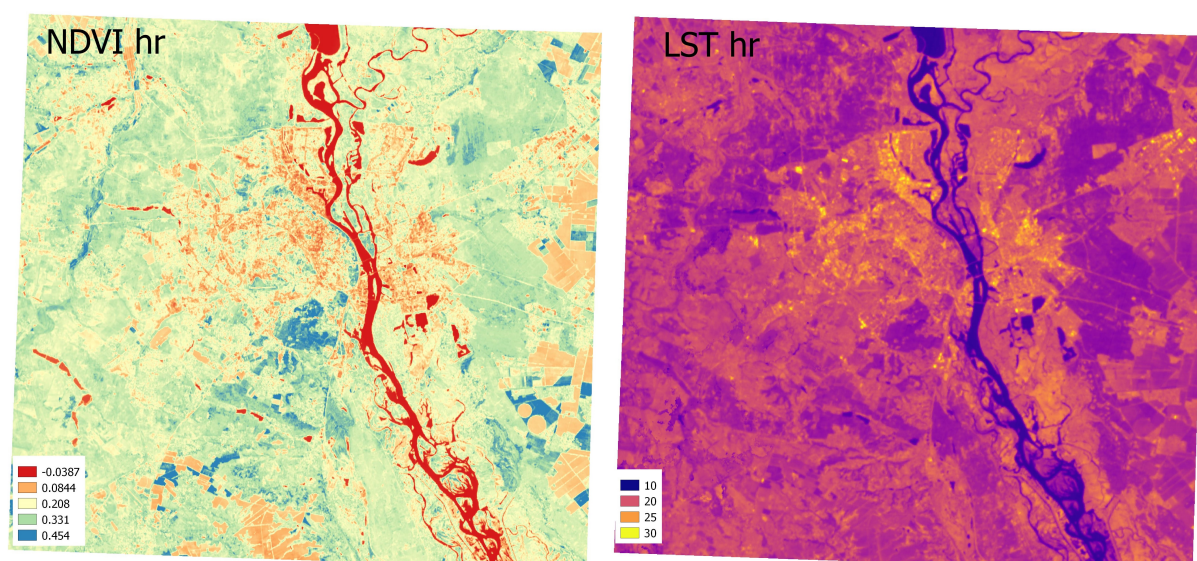


Рисунок 4.7 – NDVI та LST карти високого просторового розрізнення

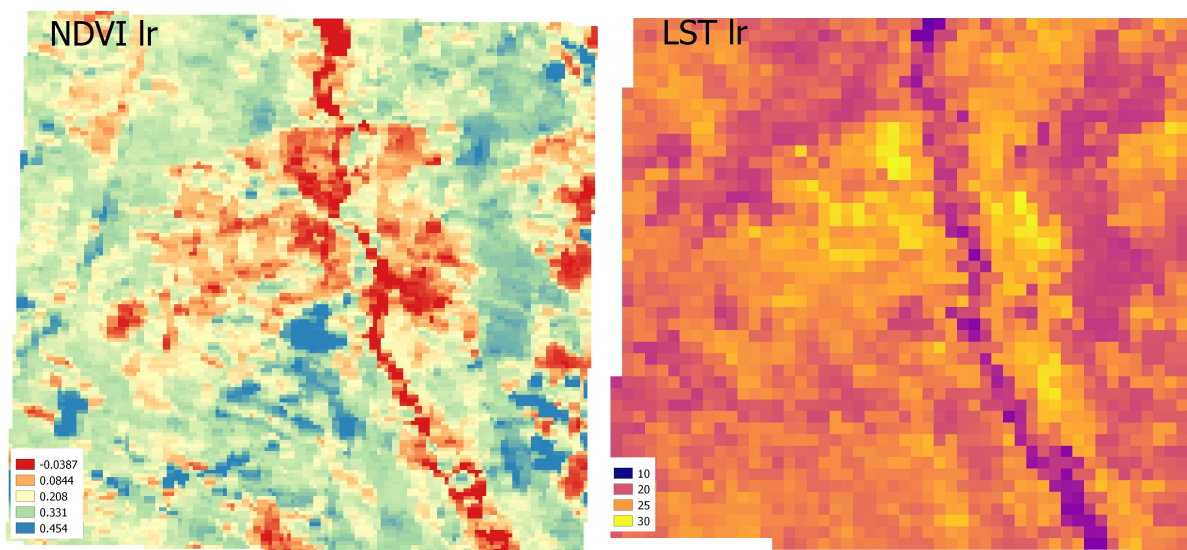


Рисунок 4.8 – NDVI та LST карти отримані за допомогою Sentinel-3

Якщо говорити про навчання нейронних мереж важливо, щоб дані на вході мали нормальний розподіл. При цьому відхилення від нормального розподілу мають пояснюватися низькою кількістю зашумлених та аномальних даних. Зрозумілим є те, що якщо датасет є зашумленим, для його використання необхідно прибрати шум. Але при цьому невелика кількість шуму навпаки сприяє збільшенню точності моделі, оскільки працює як додатковий фактор регуляризації. Часто на практиці в навчальний датасет додається невелика кількість штучного шуму. Для аналізу датасету також використовується правило, що вектор випадкових величин з нормальним розподілом є нормально розподіленою випадковою величиною. Тобто необхідно довести, що ознаки в наших випадкових векторах по окремоті мають близький до нормального розподіл. В нашому випадку наборами навчальних даних є часові ряди векторів, тобто на навчання йде 2,467,459 тензорів з розмірністю (3,4), а на тестування йдуть 2,643,986 тензорів з такою ж розмірністю. Ці тензори представляють собою пікселі високого просторового розрізнення геопросторових продуктів які відображають температуру земної поверхні та індекс вегетації у високому та низькому просторовому розрізненні. Важливим критерієм вибору території для даного експерименту є реалістичний та хороший в плані статистики розподіл класів земного покриву [57, 58, 59], в якому немає домінації одного класу, а основні класи присутні в близькій до збалансованій статистиці. На рисунку 4.9 зображено розподіл класів земного покриву для даної території.

Де перший клас це штучні об'єкти, другий клас це водні об'єкти, третій клас це дерева, а 4 клас це низькі зелені насадження (сільськогосподарські поля, луки, клумби). Для реальних даних це дуже хороший розподіл для експериментів, оскільки для більшості можливих зон для України або домінує клас 4, а всіх інших класів незначна кількість, або вони не присутні, або домінує клас 3 і всіх інших класів незначна кількість або їх немає.

Першим критерієм якості датасету є емпіричне правило 68-95-99.7

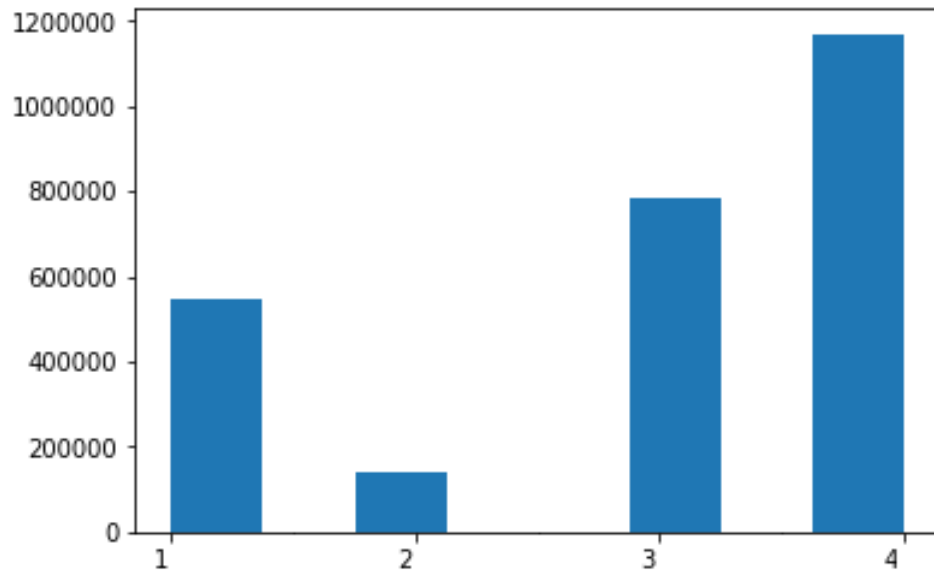


Рисунок 4.9 – Розподіл класів земного покрову для території тестування

[60]. З математичної точки зору це правило пояснюється як те, що при вимірюванні випадкової величини X з математичним сподіванням μ та дисперсією σ приймає випадкові значення. Для випадкової величини з нормальним розподілом діє правило:

$$Pr(\mu - 1\sigma \leq X \leq \mu + 1\sigma) \approx 0.68$$

$$Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 0.95$$

$$Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx 0.997$$

Також якщо відомо, що випадкова величина має мати нормальний розподіл за допомогою цього правила можна знаходити викиди, як всі елементи, що не належать до інтервалу $\mu - 3\sigma$ — $\mu + 3\sigma$, або використовуючи пов'язаний з цим правилом довірчий інтервал $\bar{X} \pm 2\frac{\sigma}{\sqrt{n}}$.

Підготовлені реальні дані проходять цей критерій. На рисунку 4.10 зображено гістограму розподілу ознак даного датасету. Середній відсоток даних, що лежать у діапазоні $\mu - 3\sigma$ — $\mu + 3\sigma$ по ознакам дорівнює 0.991, що означає, що датасет відповідає критерію та містить невелику кількість шуму.

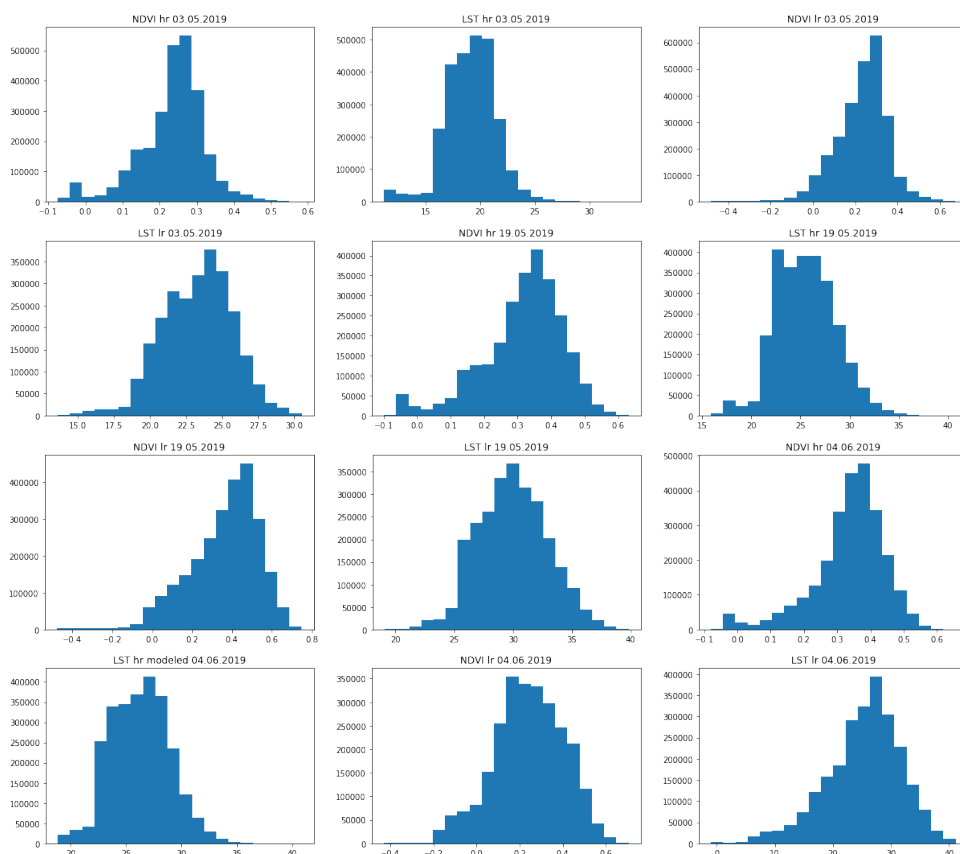


Рисунок 4.10 – Гістограми ознак вибірки, які показують розподіл даних

При цьому на рисунку 4.11 зображено гістограма значень даних для карти температури земної поверхні високого просторового розрізнення при навчанні за поточний день. Відсоток даних, що лежать між $\mu - 3\sigma$ — $\mu + 3\sigma$ дорівнює 0.994.

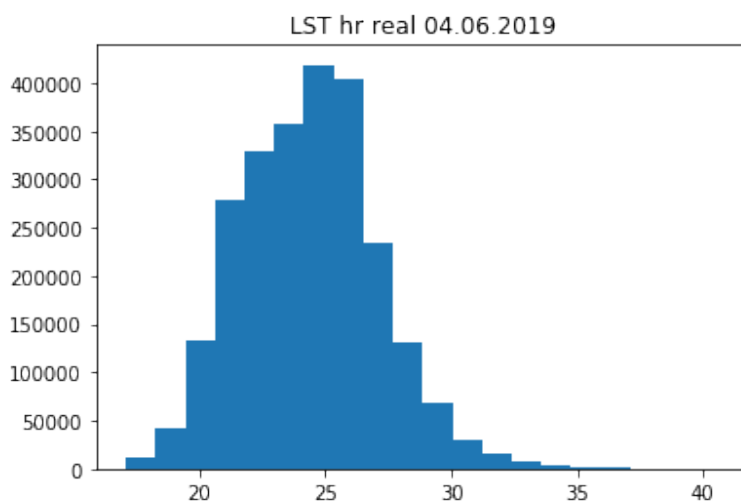


Рисунок 4.11 – Гістограми температури земної поверхні hr за 04.06.2019

Другим статистичним критерієм, використаним у аналізі датасету є тест Колмогорова-Смірнова [61], який ґрунтується на критерії узгодженості Колмогорова для перевірки двох гіпотез:

1) Гіпотеза H_1 полягає в тому, неперервна функція розподілу нашої випадкової величини F_1 дорівнює функції розподілу нормальної випадкової величини F

2) Гіпотеза H_2 полягає в тому, неперервна функція розподілу нашої випадкової величини F_1 не дорівнює функції розподілу нормальної випадкової величини F

Нехай $\rho(x) = \sqrt{n} \sup |F_n^*(y) - F_1(y)|$. Якщо гіпотеза H_1 не виконується, тоді випадкова величина X має відмінний розподіл від F_1 , який ми назвемо F_2 . За теоремою Глівенко-Кантеллі $F_n^*(y) \rightarrow F_2(y)$, для будь-якого y при $n \rightarrow \infty$. Таким чином при $F_1 \neq F_2$ існує таке значення y_0 , при якому $|F_2(y_0) - F_1(y_0)| > 0$. При цьому $|F_n^*(y) - F_1(y)| \geq |F_n^*(y_0) - F_1(y_0)| \rightarrow |F_2(y_0) - F_1(y_0)| > 0$. Помноживши цей вираз на \sqrt{n} ми отримуємо те, що $\rho(x) \rightarrow \infty$. При цьому нехай випадкова величина η має розподіл з функцією розподілу Колмогорова $K(t) = \sum_{j=-\infty}^{+\infty} (-1)^j e^{-2j^2 t^2}$, при $t > 0$. За допомогою цього розподілу можна знайти таке C за заданим ε , що $\varepsilon = P(\eta \geq C)$. Таким чином Критерій колмагорова зводиться до системи:

$$\begin{cases} H_0, \rho(X) < C \\ H_1, \rho(X) \geq C \end{cases}$$

Тест Колмогорова-Смірнова показує наскільки значущими є відмінності між двома розподілами випадкових величин. В нашому випадку, ми порівнюємо розподіл ознаки та нормального розподілу. Для цього відбувається підрахунок статистики Колмогорова-Смірнова:

$$D_{n,m} = \sup |F_n^*(y) - F_{1m}(y)|$$

Де F_n та F_{1m} імперичні функції розподілу, а n та m - це розміри вибірок для

двох розподілів. Для достатньо великого датасету гіпотеза H_0 відкидається при умові:

$$D_{n,m} > c(\alpha) \sqrt{\frac{n+m}{n * m}}$$

Значення $c(\alpha)$ можна використати підраховані для різних α , або брати як:

$$c(\alpha) = \sqrt{-\ln\left(\frac{\alpha}{2}\right) * 0.5}$$

Таким чином тест Колмогорова-Смірнова приймає вигляд:

$$D_{n,m} > \frac{1}{\sqrt{n}} \sqrt{-\ln\left(\frac{\alpha}{2} * \frac{1 + \frac{n}{m}}{2}\right)}$$

Використаний в експериментах набір даних пройшов даний тест. Для нього було обрано границю значимості $D_{n,m}$ при тестуванні гіпотези того, що даний датасет має близький для нормального розподіл, рівню 0.05. Тобто, якщо статистика $D_{n,m} > 0.05$, значить вона є значима і гіпотеза відкидається. Для нашого датасету середня статистика по ознакам дорівнює 0.044.

Третім критерієм для перевірки нормальності набору даних був вибраний критерій Шапіро-Уїлка [62], який ґрунтується на оптимальній лінійній незміщеній оцінці дисперсії до її звичайної оцінки методом максимальної правдоподібності. Статистичний критерій має вигляд:

$$W = \frac{1}{s^2} \left[\sum_{i=1}^n a_{n-i+1} (x_{n-i+1} - x_i) \right]^2$$

$$s^2 = \sum_{i=1}^n (x_i - \bar{x})^2$$

Чисельник для даного критерію є середньоквадратичним відхиленням Ллойда. Коефіцієнт a_{n-i+1} є табличним значенням, так само як і можливі критичні значення $W(\alpha)$. Тобто, якщо $W < W(\alpha)$, тоді гіпотеза про нормальний розподіл відкидається, або при рівні значимості $p > 0.05$ розподіл над яким відбувається аналіз не має суттєвих відмінностей від

нормального розподілу.

Даний критерій використовується для аналізу невеликих наборів даних, а для наборів даних, що містять більше 100 семплів він може працювати некоректно. Тому для експериментів із датасету для кожної ознаки вибиралося випадковим чином 10 семплів, для яких обраховувався критерій Шапіро-Уїлка. Дана дія повторювалася 1000 разів і отримані статистики та коефіцієнти значимості усереднювалися. В результаті набір даних пройшов даний критерій, оскільки середня статистика $W = 0.907$, а середній $p = 0.43$

Таким чином, використавши три статистичних критерії оцінки набору даних можна дійти до висновку, що набір даних має нормальний розподіл з невеликою кількістю шумів і підходить для проведення експериментів над моделями нейронних мереж.

4.4 Тестування моделі нейронної мережі для збільшення просторового розрізнення супутникових даних

Підбір гіперпараметрів для моделі нейронної мережі є експериментальною задачею і для різних територій, в залежності від розподілу класів земного покрову та розмірів території. При цьому розмір території впливає не тільки на кількість вхідних даних, але також на варіативність ознак, оскільки великі території часто відрізняються за погодними та кліматичними умовами і тому для їх коректного запам'ятовування моделлю необхідно збільшувати кількість шарів та кількість епох навчання. Тому для невеликих тестових територій можна використовувати одношарові архітектури, для території Київської області оптимальною є 19-ти шарова архітектура, а при збільшенні території необхідно також експериментувати із кількістю шарів.

Тепер порівнюємо результати, що отримуються за допомогою такого підходу із іншими класичними методами регресії на підготовленому та

проаналізованому навчальному та тестовому датасеті. Для цього було вибрано лінійну регресію, баєсову регресію, регресію на основі випадкового лісу та регресію на основі нейронної мережі багатошарового персептрону. У таблиці 1 наведено точність підрахована для різних класів земного покриву

Таблиця 4.3 – Середня помилка збільшення розрізнення температурних карт для різних типів земного покриву

Клас земного покриву	LSTM	LR	Bayes	RF	MLP
Штучні поверхні	2.28	2.48	2.48	2.4	2.39
Водні об'єкти	1.64	2.31	2.31	1.99	1.85
Деревинний покрив	2.4	2.96	2.96	2.55	2.47
Низькі зелені насадження	2.5	2.77	2.77	2.70	2.61

Як видно по точності для всіх класів LSTM модель дає стабільний високий результат точності, при цьому найближчий результат отримується за допомогою нейронної мережі багатошарового персептрону та гірший за допомогою випадкового лісу, інші моделі не наближуються до такої точності. Тобто особливість LSTM у аналізі даних із структурою часового ряду векторів є ефективною. На рисунку 3.9 та 3.10 наведено приклад температурної карти за 20.06.2019 на основі Landsat-8 та із збільшеним просторовим розрізненням, тобто результат роботи методу. Після збільшення розрізнення супутникових карт температури земної поверхні відбувається їх векторизація за допомогою підрахунку зональної статистики на полігонах міського атласу Києва. Таким чином відбувається створення інтерактивної карти, яку в подальшому можна використовувати для геопросторового аналізу. Це також викликано зручністю роботи з векторними даними у ГІС системах.

На рисунку 4.14 зображено векторну карту LST для Києва отриманої на основі зональної статистики на основі карти температури земної поверхні із збільшеним просторовим розрізненням за 07.08.2019.

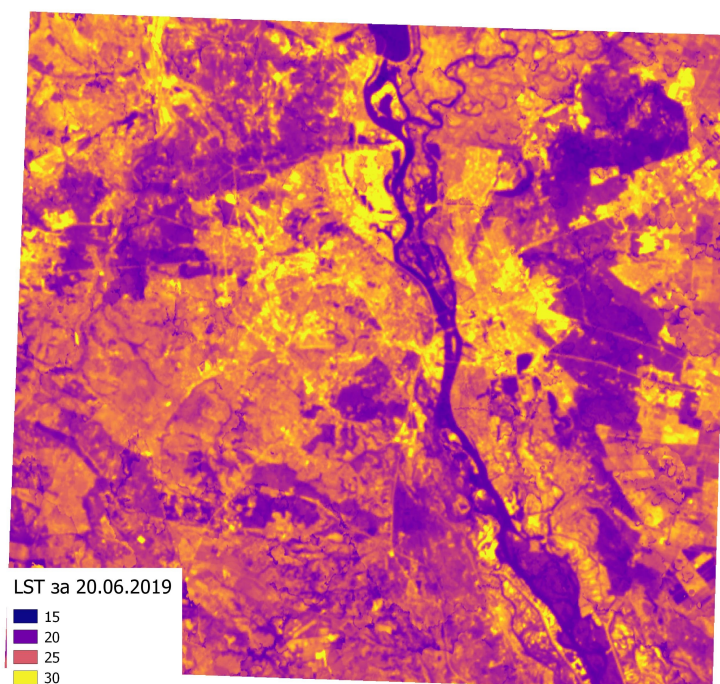


Рисунок 4.12 – Карта температури земної поверхні для міста Київ за 2019.06.20 на основі Landsat-8

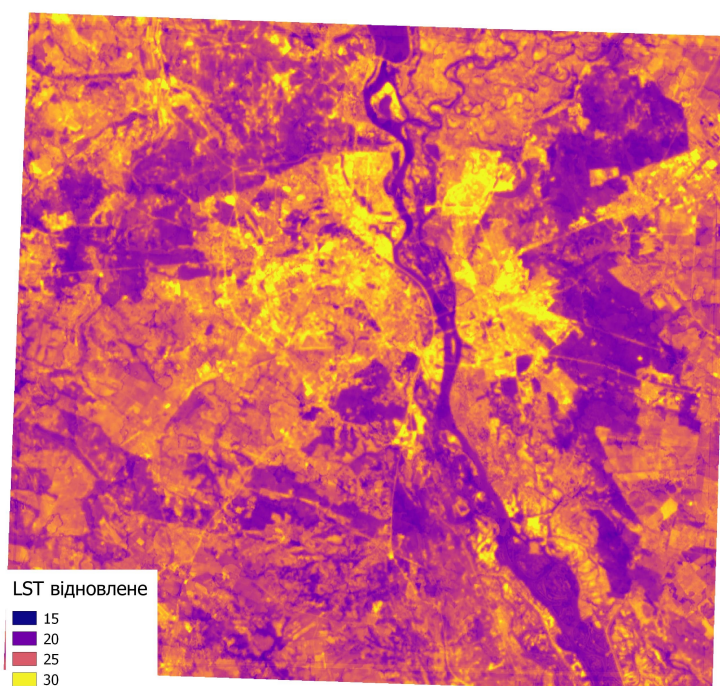


Рисунок 4.13 – Результат збільшення просторового розрізнення температури земної поверхні для міста Київ за 2019.06.20

Даний аналітичний продукт показує усереднену температуру земної поверхні для дрібних сегментів міста і дає більше можливостей для

аналізу показників урбанізації.

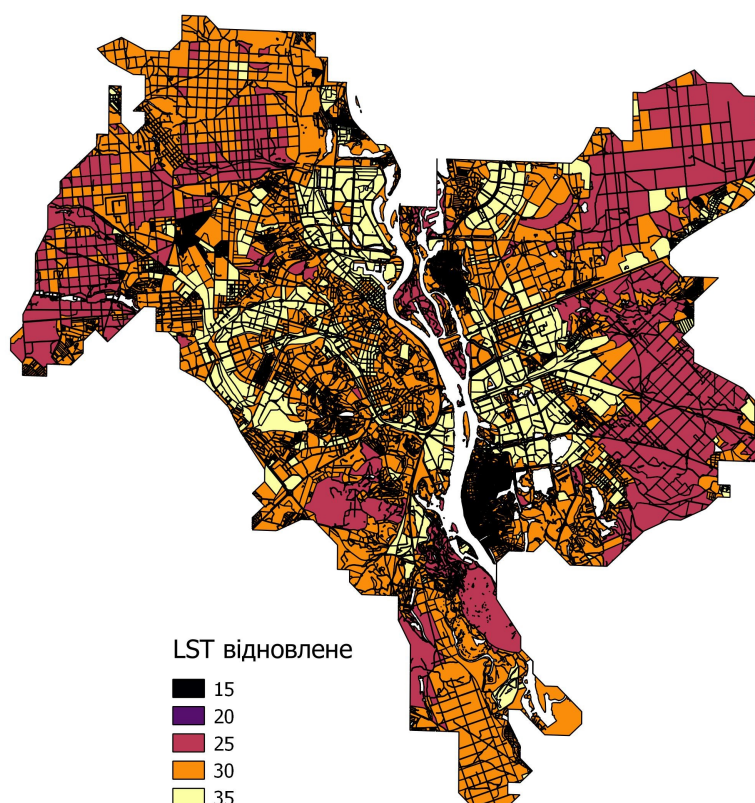


Рисунок 4.14 – Векторна карта температури земної поверхні збільшеного просторового розрізнення для міста Київ за 2019.06.20

Цей експеримент показав, що дану методологію можна використовувати для збільшення просторового розрізнення супутникових даних температури земної поверхні Sentinel-3 з високою точністю. При цьому необхідно зазначити, що навчання моделі може відбуватися як із даними з великою різницею в часі так і з маленькою, таким чином отримуючі якісні продукти температури земної поверхні з високим часовим та просторовим розрізненням.

Висновки до розділу 4

Сучасне відкрите програмне забезпечення дозволяє створювати сервіси на основі супутникових даних, що працюють в близькому до

реального часу режиму. Також вже існує велика кількість супутникових продуктів поєднання яких дає можливість створити інноваційний аналітичний продукт. В даному розділі показано як об'єднавши всі доступні дані отримані із супутників Sentinel-2, Landsat-8 та Sentinel-3 можна створити новий набір даних, який поєднує в собі високе просторове та часове розрізнення.

В даному розділі було реалізовано процедуру завантаження та обробки супутникових даних Sentinel-2, Landsat-8 та Sentinel-3. Ця процедура була використана для формування гармонізованого набору даних, який використовується в подальших експериментах.

Було розроблено нову методологію відновлення хмар на супутникових продуктах вегетаційних індексів NDVI та температурних карт. Дана методологія ґрунтується на використанні моделі машинного навчання випадкових лісів в режимі регресії і складається з двох фаз. При першій фазі модель навчається переводити зображення самого в себе для пошуку аномальних пікселів. У цьому випадку аномальними пікселями є пікселі які для свого діапазону температури мають аномально велику похибку відновлення. Аномальні пікселі інтерпретуються, як пікселі пошкоджені хмарами або тінями, які не були замасковані при використанні маски хмар. У другій фазі відбувається видалення аномальних пікселів із навчального набору даних переведення знімку самого в себе і нове навчання моделі. Після цього відбувається запуск моделі та отримання безхмарної карти NDVI та LST з відновленими хмарами.

Описано процес підготовки навчального та тестового набору даних. При цьому було проаналізовано розподіл вхідних даних для моделей за допомогою емпіричного статистичного правила 8-95-99.7, тесту Колмогорова-Смірнова та тесту Шапіро-Уїлка. Аналіз розподілу показує, що датасет є придатним для навчання та валідації нейромережевого підходу збільшення просторового розрізнення супутникових даних температури земної поверхні. На основі цих даних був проведений

експеримент по збільшенню просторового розрізнення карти температури земної поверхні для Києва та порівняння результатів отриманих результатів із іншими класичними регресійними моделями. Експерименти показали, що модель на основі LSTM дає кращий результат ніж з застосуванням класичних регресійних методів, таких як лінійна регресія, баєсова регресія, випадкові ліси та багатошаровий перцептрон.

5 ВИКОРИСТАННЯ СУПУТНИКОВИХ ПРОДУКТІВ ТЕМПЕРАТУРИ ЗЕМНОЇ ПОВЕРХНІ У ПРИКЛАДНИХ ЗАДАЧАХ

Впровадження супутникового моніторингу для території України є дуже важливим процесом розвитку цієї сфери в Україні. Супутникові технології дають можливість проводити екологічні дослідження та створювати інноваційні сервіси, які здатні вплинути на розвиток суспільства та держави. В даному розділі розглянуто прикладні задачі та дослідження, зроблені в рамках даної дисертації, які ґрунтуються на використанні карт температури земної поверхні. Дані дослідження були зроблені в рамках проєкту ERA-Planet SMURBS та пов'язані з технологіями розумних міст, моніторингу стихійних лих, оптимізації процесів планування міста та розробки програм озеленення міста.

Першою територією дослідження є місто Рівне. В ході дослідження було створено сегментацію міста по типу міського атласу. Збудовано інтерактивні векторні шари щільності забудови та щільності насаджень дерев для кожного полігону шару за допомогою зональної статистики на основі карти земного покрову. Збудовано карти температури земної поверхні з 2013 по 2018-ий рік та обраховано тренди температури земної поверхні. А також проведено аналіз супутникових показників поверхні міста на основі супутникових даних.

На основі сучасних підходів до урбаністики було розроблено методологію побудови шару рекомендацій міста на основі супутникових растрових та векторних даних. Дана методологія була протестована на місті Ірпінь.

Також на основі супутникових даних Sentinel-3 було розроблено сервіс моніторингу пожеж в Україні, а також проаналізовано пожежі за 2019-ий рік з використанням додаткової геопросторової інформації.

5.1 Аналіз характеристик поверхні міста для виявлення проблемних ділянок у місті Рівне

Велика кількість міст України зіткнулася з проблемою відсутності апарату моніторингу стану земної поверхні у місті та його екологічних показників. Дані дистанційного зондування дають можливість регулярно надавати достовірну інформацію в близькому до реального часу режимі для великої території. Цей підхід застосовується у всьому світі та в рамках проєкту ERA-Planet SMURBS був використаний для міста Рівне [63]. Дане дослідження включає аналіз стану міських дерев, температури земної поверхні міста, характеристик щільності забудови та щільності деревинного покриву та порівняння з містом Люблін. Для проведення даного дослідження були використані карти земного покриву на основі Sentinel-2 збудовані Інститутом Космічних Досліджень, вегетаційний індекс (NDVI), супутникові знімки Landsat-8, що використовуються для побудови карт температури земної поверхні та підрахунку трендів вегетації та температури земної поверхні[64].

Для проведення якісного аналізу отриманих геопросторових даних необхідно перейти від растрової до векторної форми, тобто необхідно провести сегментацію міста на досить однорідні зони за типом використання земель по типу Urban Atlas. Для цього використовуючи методологію сегментації міст для побудови векторних карт по типу Urban Atlas, були використані дані Open Street Map, та карти класифікації земного покриву згідно з наступними правилами: полігон можна розділити на кілька полігонів за допомогою наземних транспортних вузлів, малі полігони можуть поєднуватися один з одним за умови однакової щільності забудови та типу використання землі, мінімальний розмір багатокутників не може бути менше 100 * 100 метрів. Використовуючи такий підхід, формування векторної сегментації міст та зональної статистики на основі супутникових продуктів за полігонами. В

результаті таких дій отримується векторний шар геопросторових даних, який дає більше можливостей для проведення геопросторового аналізу та візуалізації даних, ніж растровий підхід.

Першим кроком аналізу є підготовка векторних слоїв щільності забудови (рис. 5.1) та щільності деревинного покриву (рис. 5.2) на основі карт класифікації земного покриву.

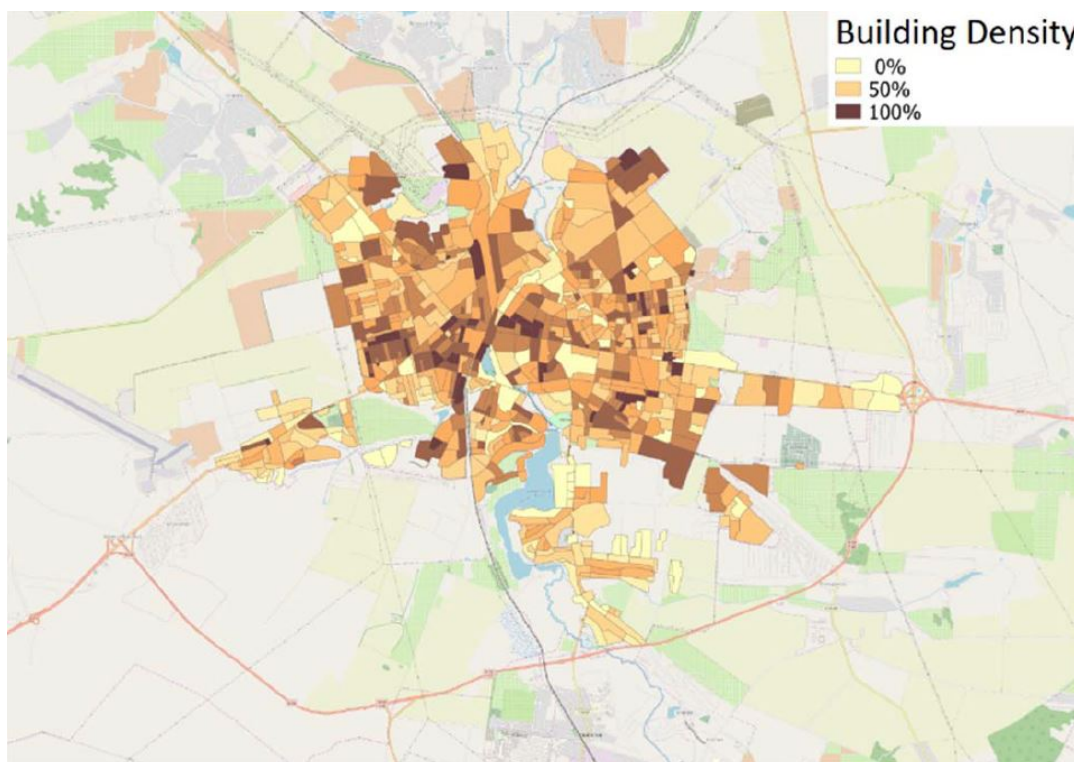


Рисунок 5.1 – Шар щільності забудови для міста Рівне на основі векторизації по типу міського атласу

За допомогою цих векторів обчислюється розподіл зон відповідно до щільності забудови та щільності дерев (таблиця 5.1). Розподіл показує дуже високу щільність забудови та дуже малу кількість зон з високою щільністю дерев. 92% усіх полігонів мають щільність дерев менше 30%. Також 85% полігонів мають щільність дерева менше 10%. Порівняно з європейськими містами, це дуже поганий показник. Наприклад, у Любліні лише 57% полігонів міського атласу мають щільність дерев менше 10 відсотків. У той же час у Рівному лише 2% полігонів мають

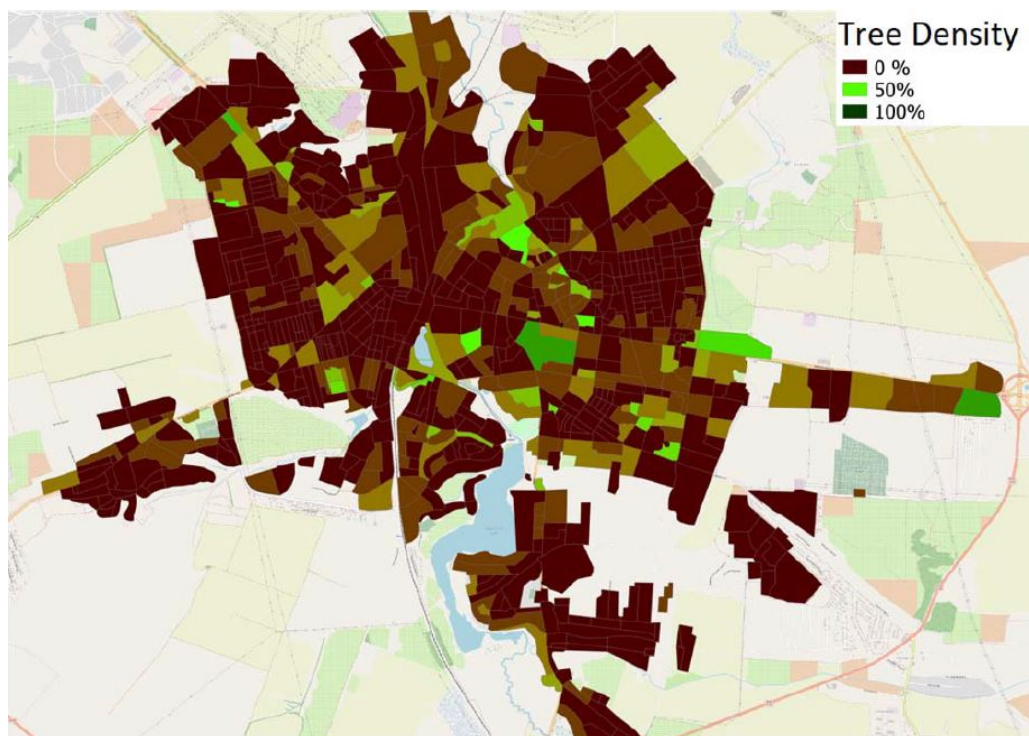


Рисунок 5.2 – Шар щільності дерев для міста Рівне на основі векторизації по типу міського атласу

понад 50% щільності дерев, коли для Любліна це 10% полігонів.

Таблиця 5.1 – Розподіл кількості полігонів за щільністю забудови та щільністю міських дерев

Щільність	Забудова	Дерева
Щільність <30%	10%	90%
30% <= Щільність <50%	16%	6%
50% <= Щільність <80%	51%	2%
80% <= Щільність <90%	21%	1%
Щільність >= 90%	2%	1%

Наступним кроком є аналіз температурних показників міста. Для цього було збудовано карти середньої літньої температури земної поверхні для 2013-2018 років. Середній показник LST для Рівного у 2018 році (рис. 5.3) дорівнює 26 градусів Цельсія. Використовуючи векторизовану класифікаційну карту та карту LST (рис. 5.4), було знайдено зони, які

мають щільність забудови понад 90% та LST більше 27 градусів Цельсія. Для пониження температури земної поверхні в цих зонах рекомендовано проводити озеленення, оскільки тренд температури поверхні землі протягом 5 років вказує на подальше підвищення LST. У деяких зонах за 5 років середня температура поверхні землі зросла на 2 градуси. Проаналізувавши отримані результати, було також виявлено зони, в яких температура поверхні землі, незважаючи на досить щільну забудову, є дуже низькою для міста. Такі ділянки характеризуються дотриманням правил посадки зелених насаджень та правильним доглядом за рослинами. Інші виявлені випадки мають низьку забудову та високу LST. Ці зони мають низький рівень NDVI. Це видно на наземних знімках, в літній час на цих зонах вигорає трава та опадає листя з дерев.

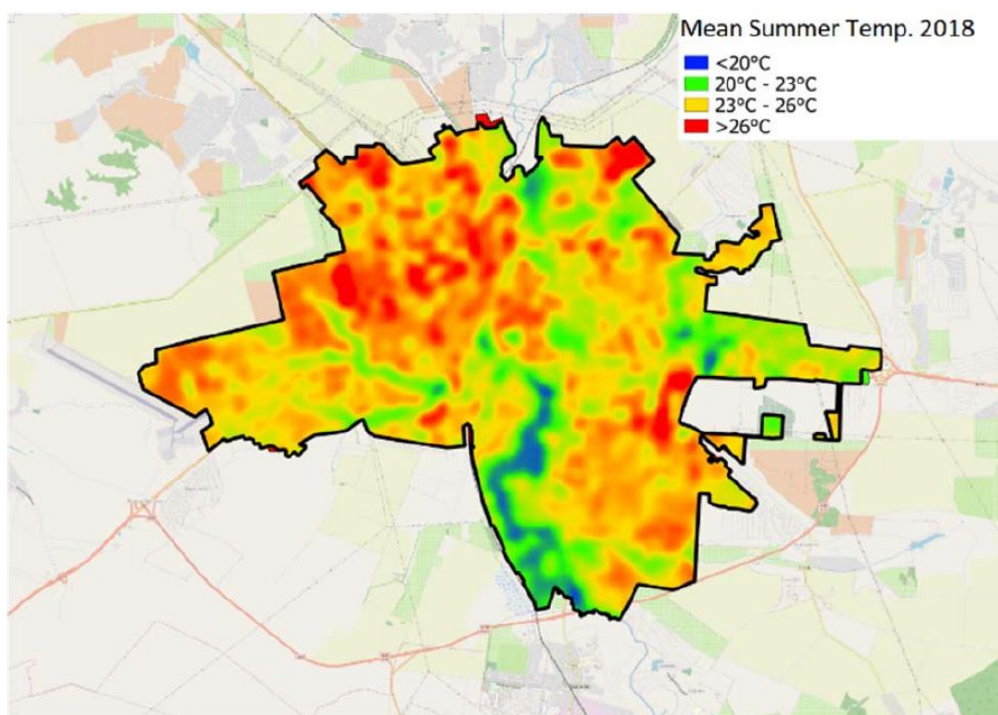


Рисунок 5.3 – Середня літня температура земної поверхні за літо 2018-го року у місті Рівне

Залежність температури земної поверхні від кількості дерев у полігонах можна помітити у таблиці 5.2. В ній продемонстровано середня температура поверхні землі для полігонів які мають відповідну щільність

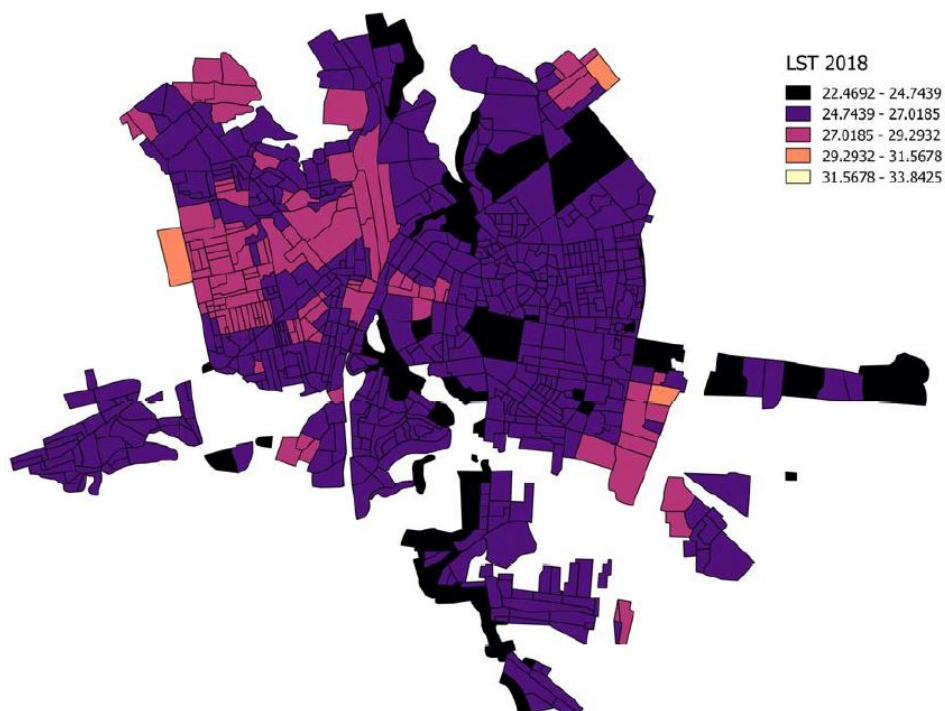


Рисунок 5.4 – Середня літня температура земної поверхні за літо 2018-го року у місті Рівне у векторному вигляді

забудови та щільності деревинного покриву.

Даних за 6 років є достатньо, щоб збудувати лінійну регресійну модель та продовжити тренд температури земної поверхні для Рівного, таким чином можна оцінити довгострокові наслідки активної урбанізації міста при відсутності змін у алгоритмах керування містом та організації

Таблиця 5.2 – Середня літня температура поверхні землі для полігонів Рівного в залежності від щільності забудови та деревинного покриву

Щільність	Забудова	Дерева
Щільність <30%	26.14	26.97
30% <= Щільність <50%	26.53	26.95
50% <= Щільність <80%	27.05	26.22
80% <= Щільність <90%	27.29	25.63
Щільність >= 90%	27.21	25.07

планування міста. На рисунку 5.5 показано отриману таким чином температуру земної поверхні за 2023 рік.

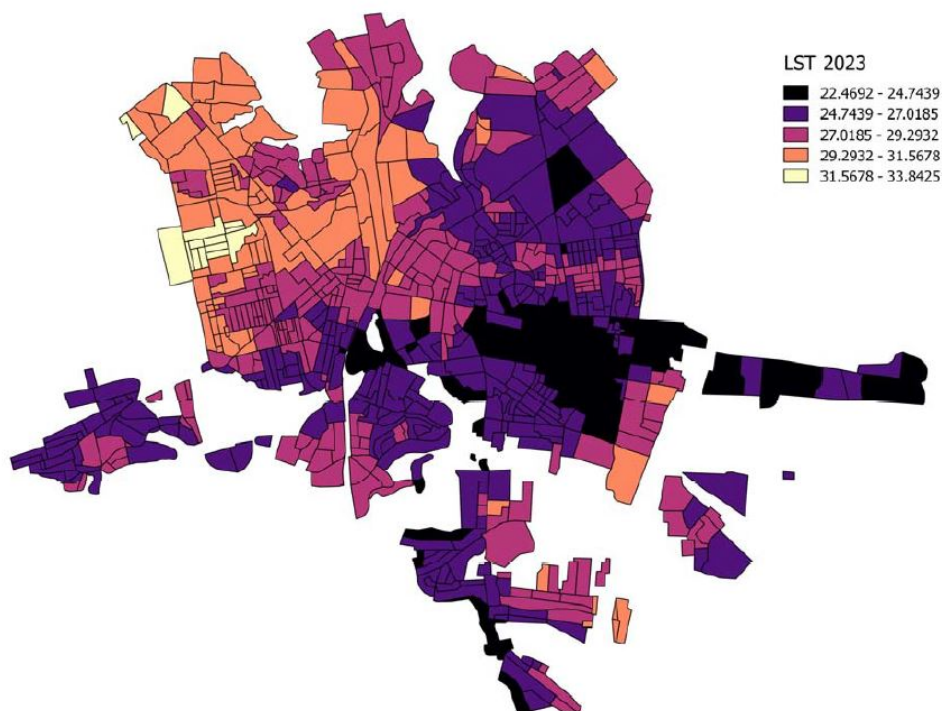


Рисунок 5.5 – Продовження тренду температури земної поверхні до 2023-го року

Останнім кроком є побудова векторизованих карт вегетації NDVI збудованої по масці міських дерев карти класифікації земного покрову (рис. 5.6). За допомогою цього продукту були виявлені зони з дерева в поганому стані. Основна проблема - неправильна обрізка дерев, яке заважає рослині продовжувати свій розвиток, а також часто спричиняє загибель рослини. Така обрізка дерев у місті не дозволяє їм виконувати свою роль міських дерев, що відображено на температурній карті.

Було також побудовано статистичний прогноз NDVI на 2023 (рис. 5.7) рік за тією ж методологією, що була використана для LST, яка показує як може змінитися стан дерев при відсутності змін у стратегії озеленення.

Таким чином було отримано аналітику показників поверхні міста для міста Рівне, яке було надано ГО "Сад Історій Рівного" [65] та доступні в

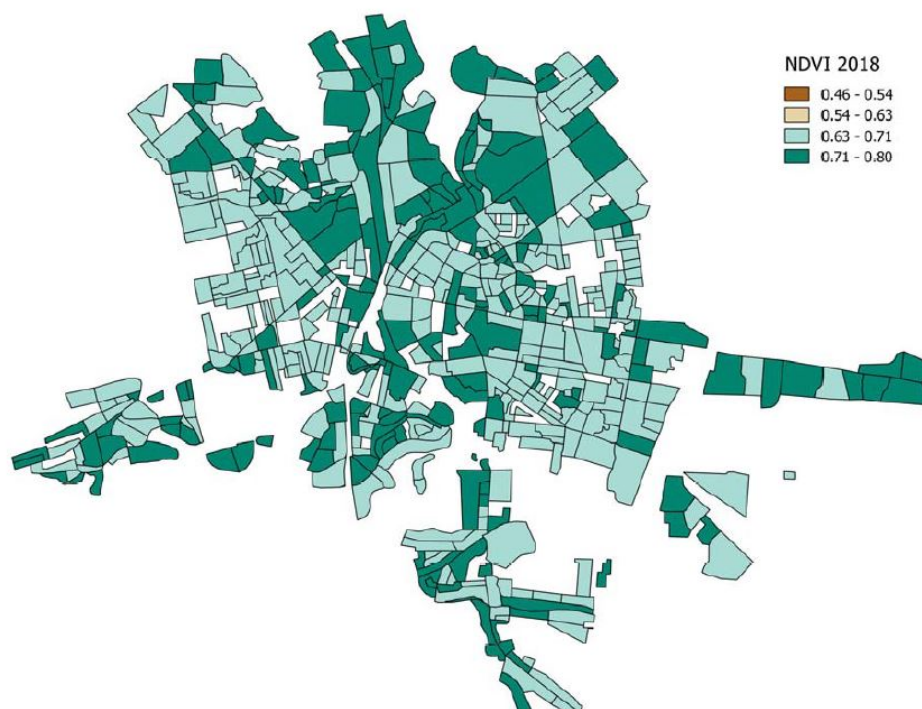


Рисунок 5.6 – NDVI міських дерев за 2018-ий рік у векторній формі



Рисунок 5.7 – NDVI міських дерев при продовженні тренду зміни
вегетації за 2023-ий рік у векторній формі

публікації, на основі якої складається програма озеленення міста Рівне, яку розробляє ГО "Сад Історій Рівного". А також отримані дані доступні на офіційному геопорталі відкритих даних Управління містобудування та архітектури виконавчого комітету Рівненської міської ради. [66].

5.2 Використання супутникових продуктів для розробки плану озеленення міста

Карти LST виступають індикатором проблеми містобудування. Висока температура поверхні землі є причиною поганої якості ґрунтів, води та повітря у міській місцевості, тому її слід контролювати та прогнозувати під час планування містобудівних процесів. При виявленні високої температури поверхні земель у зонах міста та позитивної тенденції щодо її зростання, стратегія розвитку міста у цих зонах має бути змінена з урахуванням цих показників.

Карти NDVI виступають індикатором стану міських дерев. Завдяки NDVI показано, як інтенсивно живі рослини активно здійснюють процес фотосинтезу. Низькі значення NDVI, отримані для пікселів у масці міських дерев, можна трактувати як рослинність у поганому або хворому стані, тоді як високі значення NDVI - як рослинність у хорошому стані за умови однакової стадії вегетаційного періоду для рослин над якими проводиться аналіз, а також низькому відсотку змішаних пікселів. Пікселі які є мішаними, тобто у них присутні більше ніж 1 клас земного покрову мають викидатися із аналізу.

Обчислюючи середні значення щільності дерев, щільності забудованої площі, LST та тенденції LST для полігонів Urban Atlas, можна отримати аналітичні продукти для планування стратегії озеленення міста. Для надання рекомендацій щодо подальших дій пропонуємо використовувати класифікацію 4 класів отриманих на основі дерева рішень, побудованого на характеристиках полігонів. Важливо

розділити значення температури земної поверхні для того щоб встановити нормальне значення температури земної поверхні для зон міста. Це можна зробити за допомогою аналізу середніх значень LST та тенденції LST для цих полігонів. Після цього можна проводити класифікацію за наступними правилами:

– Перший клас - зони з найвищими пріоритетами для планування змін, і ці зміни можна здійснити за допомогою стратегії озеленення. Цей клас може бути отриманий за логічними правилами: LST вище середнього LST, тенденція LST вище, ніж середній тренд LST, щільність забудови вища 30 відсотків, яка може бути визначена як нормальна міська поверхня (полігони належать до класів Urban Atlas між 11220 : Міська поверхня середньої щільності забудови (SL: 30% - 50%) та 11100: безперервна міська забудова (SL > 80%)), а сума щільності дерев та забудованої площі не дорівнює 100 (на ділянці є можливість саджати дерева). Тож інтерпретація цього класу - це територія з проблемними тенденціями LST та високим LST, на яких можливо покращити ситуацію, використовуючи стратегію озеленення, тому можна рекомендувати зробити висадку дерев.

– Другий клас - зони з найвищими пріоритетами для планування змін, і ці зміни неможливо здійснити за допомогою стратегії озеленення. Цей клас може бути отриманий за логічними правилами: LST вище середнього LST, тенденція LST вище середнього показника, належить до нормальної міської поверхні, а сума відсотків дерев та забудованої площі дорівнює 100 (на ділянці немає місця без наявного місце для нових дерев). Тож інтерпретація цього класу - це область із проблемними тенденціями LST та високим LST, на яких неможливо покращити ситуацію за допомогою класичної стратегії озеленення. Щоб покращити ситуацію, необхідно істотно змінити земельний покрив у деяких зонах від забудованої території до рослинної зони, змінивши покриття штучних об'єктів, створенням вертикальних садів або охолодженням поверхні іншими можливими способами. Також для покращення показників таких

зон необхідно використовувати карти NDVI міських дерев за допомогою якої можна виявити проблемні дерева для яких потрібно провести покращення їхніх умов догляду або заміну.

– Третій клас - зони з другими пріоритетами планування змін. Це зони, які належать до нормальної міської поверхні і LST нижче середнього значення LST, і на яких можна садити дерева або зони, які не належать до нормальної міської тканини і мають тенденцію LST вище ніж середня тенденцію LST. Для цієї зони можна використовувати аналогічну стратегію, як для першого класу, але зміни на них мають не такі високі пріоритети.

– Четвертий клас - зони, в яких ситуація хороша з точки зору наявних показників. Це зони, які належать до нормальної міської поверхні та LST нижче середнього значення LST як і тенденція LST або зони, які не належать до нормальної міської поверхні і мають тенденцію LST нижче, ніж середня тенденцію LST. Ці зони можна трактувати як стійкі зони.

5.3 Створення рекомендацій для програми озеленення міста Ірпінь

Міський Атлас міста Ірпінь, збудований Інститутом Космічних Досліджень дає можливість надати геопросторовий аналіз особливостей міської території, таких як щільність дерев (рис. 5.8) та щільність забудови (рис. 5.9). Ці функції можна використовувати для визначення зон із завершеною стадією розвитку та моніторингу розвитку міста. Також з допомогою цих даних можливо визначити оптимальну стратегію озеленення міста. Наступним типом продукції є карти температури земної поверхні. Використовуючи тепловий інфрачервоний діапазон Landsat-8, було побудовано карту середньої літньої температури земної поверхні для міста 2019-го року. Для кожного зображення пікселі,

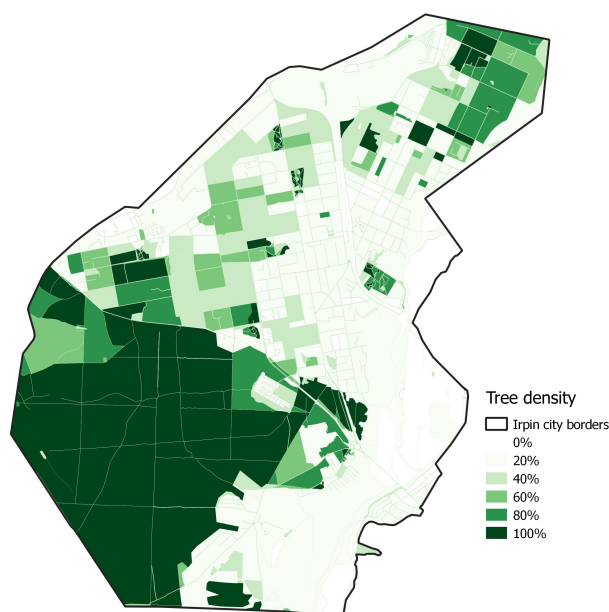


Рисунок 5.8 – Шар щільності дерев для міста Ірпінь на основі міського атласу

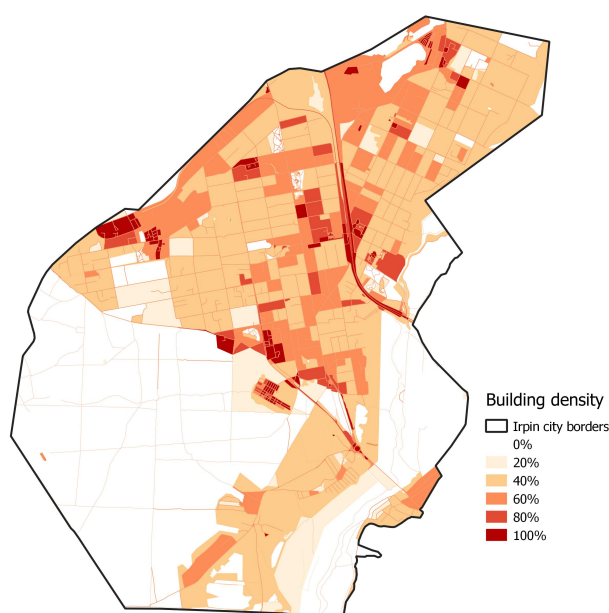


Рисунок 5.9 – Шар щільності забудови для міста Ірпінь на основі міського атласу

пошкоджені хмарами, відновлювались за допомогою регресійної моделі на основі маски хмар, карти наземного покриття, часових рядів NDVI та часових рядів LST для попередніх зображень. Більше детально процедура відновлення хмар описана в розділі 3.2. На основі отриманих карт LST за

літній період була отримана карта середньої літньої температури земної поверхні у растровій формі (рис. 5.10) та векторній формі (рис. 5.11).

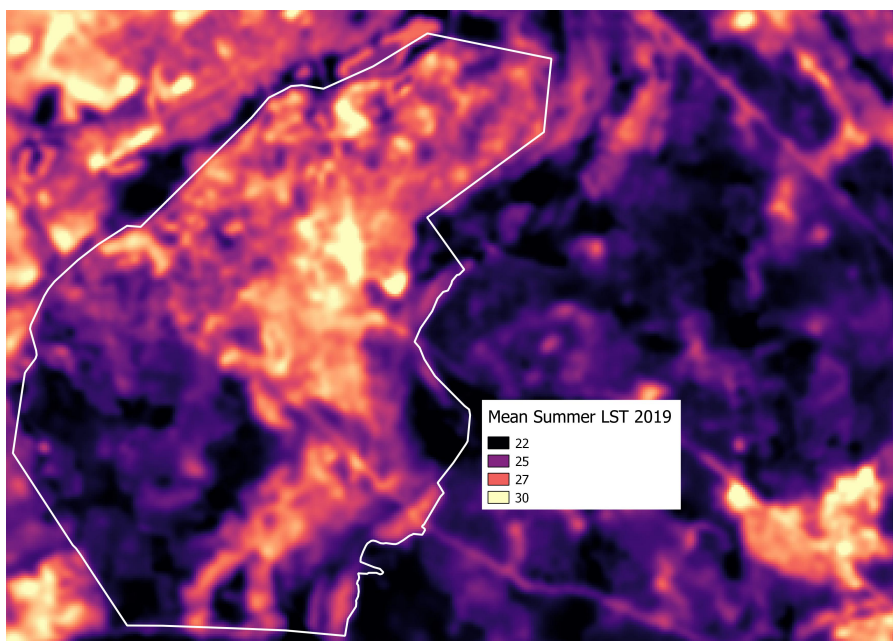


Рисунок 5.10 – Середня Літня Температура Земної Поверхні для міста Ірпінь 2019-го року у растровій формі

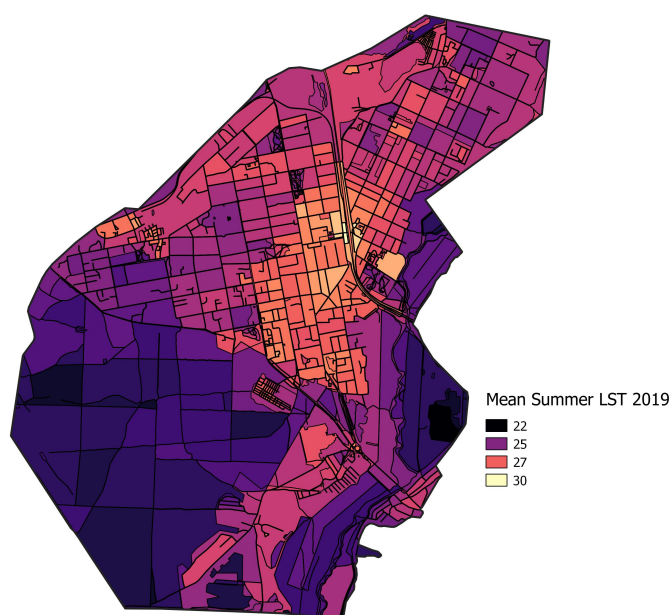


Рисунок 5.11 – Середня Літня Температура Земної Поверхні для міста Ірпінь 2019-го року у векторній формі

Використовуючи міський атлас міста Ірпінь, карти щільності забудови, карти середньої літньої температури поверхні землі та тенденції середньої літньої температури поверхні землі протягом принаймні 5 років, можна побудувати рекомендації за методикою описаній у розділі 5.2 (рис. 5.12.). Шар рекомендацій планування стратегії озеленення вказує на 82 зони

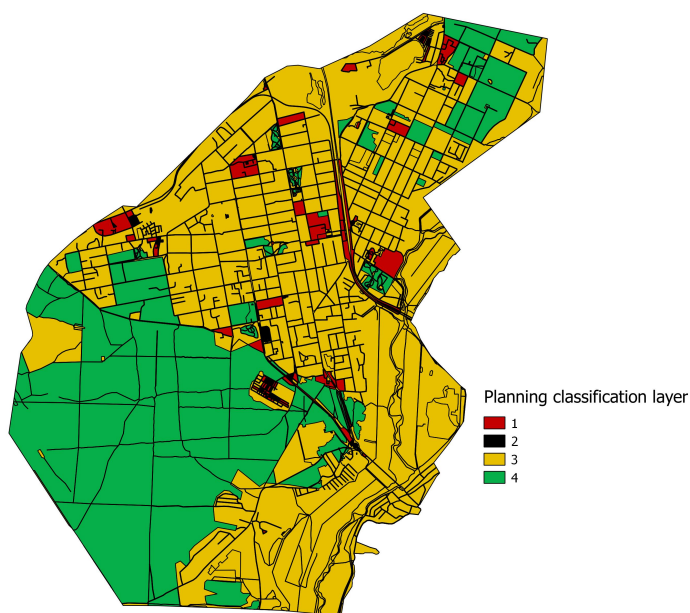


Рисунок 5.12 – Шар рекомендацій для програми озеленення

першого класу. Проблеми високих значень LST та тренду LST можуть бути вирішені шляхом збільшення кількості висаджених дерев або будь-якої іншої вегетації. Оскільки було класифіковано 85 зон другого класу, для цієї зони рішення повинно знаходитись з високим пріоритетом, оскільки тенденція LST для них вище, ніж для будь-яких інших класів. При цьому варто зазначати, що не обов'язково саджати рослинність у відкритий ґрунт, можна використовувати декоративні горщики та інші способи озеленення. Проте висадка у відкритий ґрунт найбільш ефективна, оскільки таким чином ґрунт матиме доступ до збільшення вологості через опади та збільшить швидкість охолодження, зменшивши ефект міського термального острова. 579 міських зон були віднесені до третього класу, ці зони в 2019 році не мають критичних проблем, але їх

потрібно покращувати також. Через велику кількість зелених зон у місті, 216 зон були віднесені до четвертого класу. Проблеми 1-3 класів можна вирішити також шляхом обробки міських дерев, модифікації технології догляду за вуличними деревами та їх заміною. Для прийняття цього рішення можна використовувати карти NDVI міських дерев з просторовою роздільною здатністю 10 м (рис. 5.13) та тенденцією NDVI (рис. 5.14). Ці карти показують стан дерев, та на основі них можна вибрати одне з трьох перерахованих рішень для проблемних ділянок.

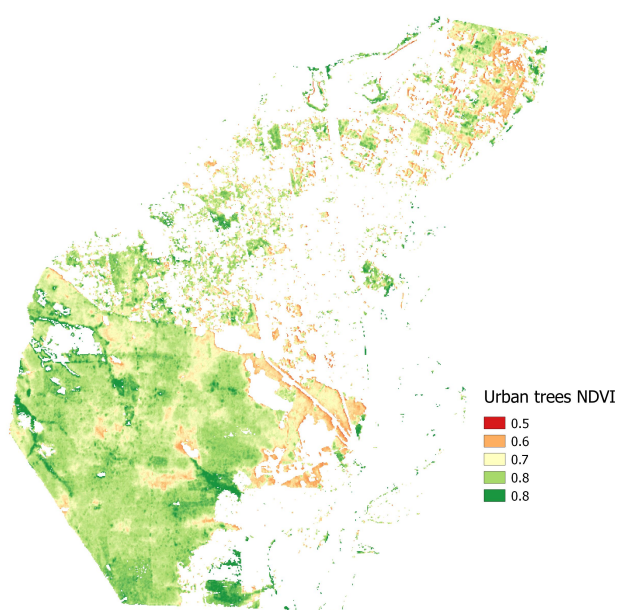


Рисунок 5.13 – NDVI міських дерев у 2019-ому році

Як видно на показниках NDVI міських дерев у 2019-ому році погіршується стан дерев у лісовій зоні, що знаходиться на південному заході міста та стан дерев безпосередньо біля щільно забудованих зон міста. Це вказує на деградацію ґрунтів у цих зонах, оскільки тренд зміни вегетації на землі є суб-індикатором індикатору Цілі Сталого Розвитку 15.3.1, який використовується для визначення деградованих земель за методологією ООН [68]. Для побудови таких аналітичних продуктів міських дерев важливо використовувати часовий ряд карт земного покриву, оскільки вирубки дерев по факту спричиняють різке

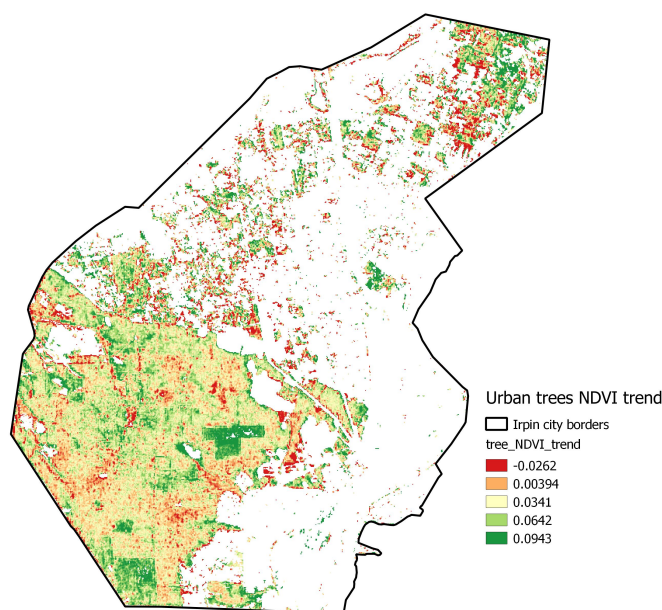


Рисунок 5.14 – тренд NDVI міських дерев за 7 років

зменшення показників NDVI і навіть при посадці нових дерев для вирівняння тренду необхідно декілька років. Тому молоді дерева не мають бути включені в маску дерев. Мішані пікселі можуть бути визначені за допомогою супутникових даних за допомогою аналізу графіків вегетаційних індексів за кілька років. Для цього необхідно вирахувати середнє значення вегетаційного індексу на початку вегетаційного сезону та порівняти його значення із середнім значенням для дерев у цей період часу. Якщо вегетаційний індекс є відносно невеликим, це означає що це змішаний піксель.

5.4 Виявлення пожеж за допомогою карт температури земної поверхні

Однією із головних задач яку на сьогоднішній день розв'язують супутникові продукти температури земної поверхні є виявлення та моніторинг активного вогню. Завдяки високому часовому розрізненню супутникових даних та великого покриття є можливість створення глобальних систем виявлення вогню на рівні країн. Для України дана

проблема є досить гострою, через велику кількість болотяних та торф'яних ґрунтів, які мають високу пожежонебезпеку. На рисунку (5.15) зображено території з такими типами ґрунтів для України на основі офіційної кадастрової карти України.

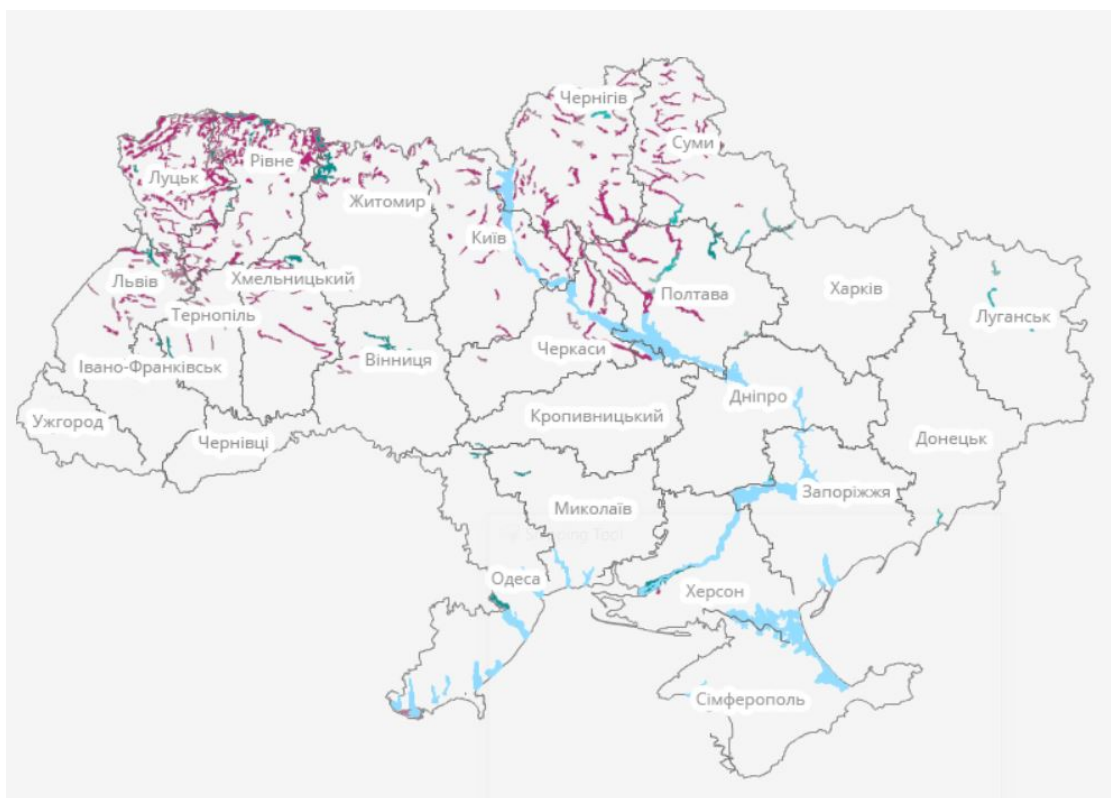


Рисунок 5.15 – карта торф'яних та болотистих ґрунтів України
представлена на інтерактивній карті

<https://superagronom.com/karty/karta-gruntiv-ukrainy>

Особливістю торфових пожеж є те що часто вони проходять під землею, без відкритого вогню. Таким чином багато оптичних способів виявлення торфових пожеж не є дієвими. При цьому термальні карти є найкращим джерелом інформації для вирішення цієї проблеми, оскільки на них торфові пожежі відображаються як аномально висока температура земної поверхні. При цьому ще однією серйозною проблемою боротьби із пожежами на торфовищах є те що, велика кількість з них розташовані у місцях з поганою інфраструктурою, що спричиняє

складнощі для доставки води та гасіння. При цьому ситуація ускладнюється через велику швидкість розповсюдження підземного вогню. Саме тому найчастіше боротьба з торфовими пожежами починається із зупинки розповсюдження за допомогою рвів. Для того щоб ефективно проводити дані роботи необхідно коректно оцінити зону де відбувається пожежа, а також оцінити швидкість розповсюдження вогню. Для цього необхідно використовувати супутникові дані з високим просторовим розрізненням і високим часовим розрізненням. Дані Landsat-8 є хорошим джерелом інформації для обрахунку площі зони ураженої вогнем, але часове розрізнення в 16 днів не дає можливості проводити моніторинг пожеж у близькому до реального часу режимі. Дані Sentinel-3 в свою чергу дають можливість з високою точністю визначити місцезнаходження пожежі і проводити моніторинг пожеж у режимі близькому до реального часу, проте просторове розрізнення 1 кілометр не дає можливості коректно оцінити площу враженої вогнем поверхні і швидкість розповсюдження вогню.

Саме тому для розв'язання задачі виявлення пожеж в близькому до реального часу на рівні країни були обрані дані Sentinel-3. Стабільність роботи сервісу забезпечується за рахунок того, що це нова сім'я супутників, місія Sentinel-3 почалася у 2016-ому році, а очікуваний час життя складає 20 років. Також методологія активного виявлення пожежі проста для автоматизації, На рисунку (5.16) наведено інтерфейс сервісу виявлення вогню на основі супутникових даних Active Fire in Ukraine розробленому в проєкті ERA-Planet SMURBS.

Для побудови карт активного вогню на основі даних Sentinel-3 була використана модифікація алгоритму, описаного Giglio [67]. Ця методологія використовувала діапазони температурну яскравість для продуктів Sentinel-3 рівня 1 RBT (випроміннювання TOA), виміряні радіометром температури моря та поверхні суші (SLSTR). Для визначеної пожежі в денний час використовується формула:

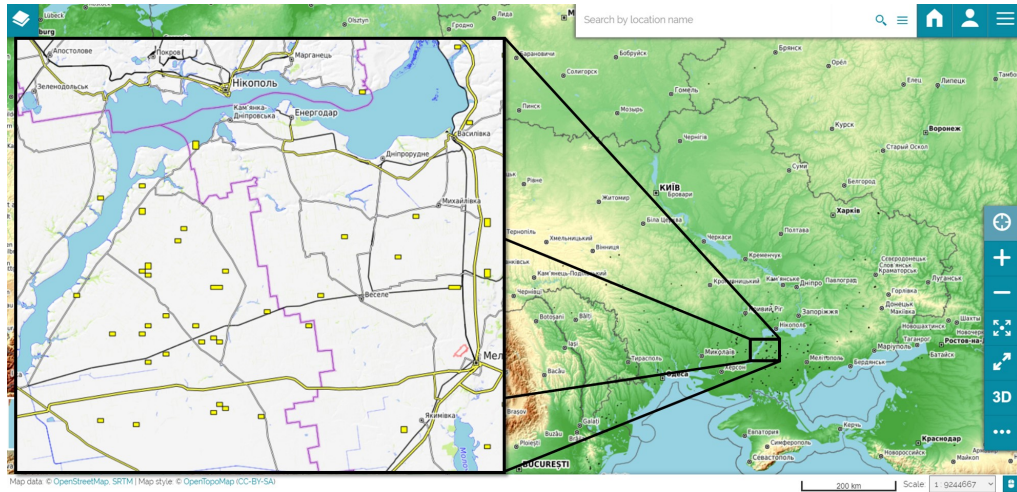


Рисунок 5.16 – Інтерфейс сервісу Active Fire

$$F1_BT_in > 325 \cap (F1_BT_in - F2_BT_in) > 18$$

Де $F1_BT_in$ та $F2_BT_in$ це температурна яскравість $F1$ та $F2$ отримана сенсором у вигляді інтенсивності випромінювання в двох температурних діапазонах після проходження через атмосферу.

Для нічних знімків формула є подібною, проте різниця температур в різних температурних діапазонах є різною:

$$F1_BT_in > 325 \cap (F1_BT_in - F2_BT_in) > 15$$

Наступним кроком обробки даних є перетин з картою класифікації земного покриву і отримання інтелектуальних аналітичних продуктів, таких як карти пожеж на торфовищах в поєднанні із картою ґрунтів і картою паління стерні через перетин із маскою посівів.

Всього за період Літо-Осінь 2019-го року даний сервіс виявив 4715 різних місць охоплених вогнем, при цьому 2804 з них це паління стерні на сільськогосподарських ділянках поблизу населених пунктів. На рисунку (5.17) продемонстровано карту пожеж на сільськогосподарських полях для України за цей період. Це є суттєвою проблемою для України, оскільки такі пожежі призводять до серйозних екологічних проблем. По перше паління стерні на полях призводить до згорання верхнього

продуктивного шару гумусу, що в свою чергу спричиняє пониження продуктивності сільськогосподарських ґрунтів та їх опустелювання [68]. По друге відбувається серйозне забруднення повітря [69, 70]. Будь які пожежі можуть спричинити більші пожежі з якими складно боротися. Ця проблема особливо актуальна у регіонах з великою кількістю болотистих та торф'яних ґрунтів, у яких пожежі на полях можуть спричинити підземні пожежі торфу. За картами класифікації земного покриву видно, що найчастіше спалювання стерні відбувається для полів зернових культур та полів кукурудзи. При цьому найбільше пожеж на сільськогосподарських полях зафіксовано в Миколаївській та Херсонській областях. Для цих областей зернові культури та кукурудза є мажоритарними культурами.

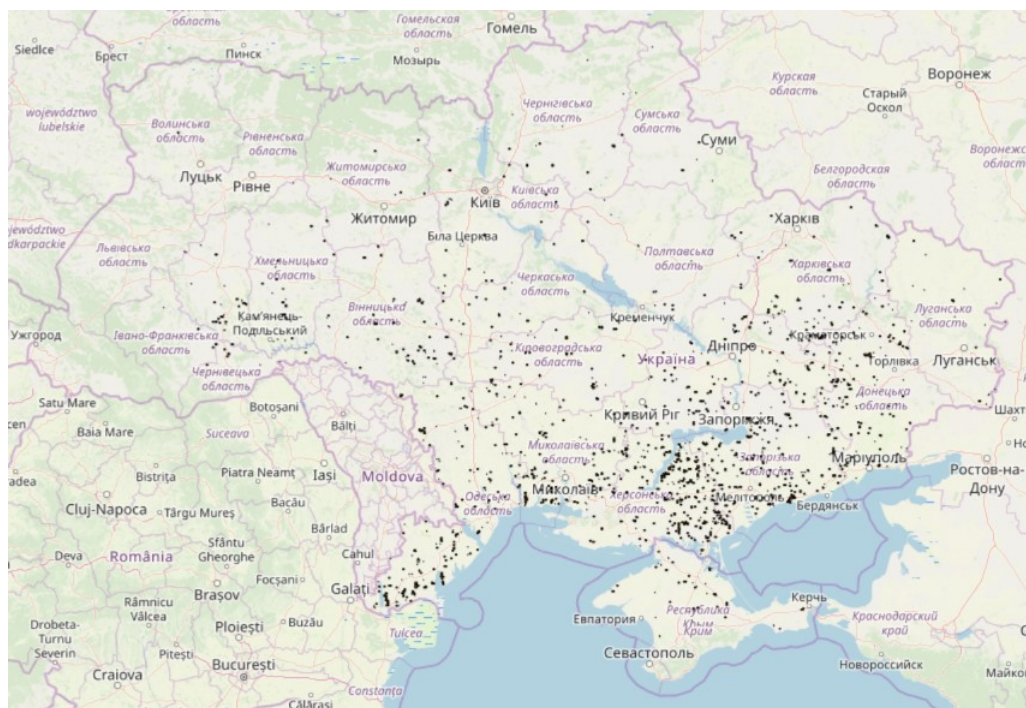


Рисунок 5.17 – Пожежі на сільськогосподарських полях для території України

Висновки до розділу 5

В рамках проєкту ERA-Planet SMURBS було проведено дослідження характеристик поверхні землі для міста Рівне. В ході даного дослідження було створено інтерактивну карту сегментації міста по типу міського атласу для подальшого геопросторового аналізу. На основі карт класифікації земного покриву було отримано шари щільності міської забудови та міських дерев. На основі супутникових даних Landsat-8 було створено карти температури земної поверхні та тренду температури, а також за допомогою об'єднання даних Landsat-8 та Sentinel-2 було збудовано карти NDVI міських дерев та їх тренди. Об'єднавши отримані інтелектуальні продукти було проведено аналіз розподілів типів поверхонь, залежність температури земної поверхні від щільності забудови і щільності дерев, а також проведено порівняння міста Рівне з містом Люблін. Даний аналіз показав проблемні зони міста в яких необхідно в першу чергу проводити програму озеленення. На основі даного дослідження ГО "Сад Історій Рівного" сформував документ "Рівне нагрівається: Як уникнути аномальних спек завдяки озелененню" на основі якого вони проводять розробку програми озеленення міста Рівне. Отримані дані під час дослідження опубліковані на офіційному геопорталі відкритих даних Управління містобудування та архітектури виконавчого комітету Рівненської міської ради.

На основі сучасних практик урбаністики було розроблено методологію побудови шару рекомендацій для програм озеленення міст на основі супутникових даних. Дана методологія поєднує супутникові дані у векторному вигляді та класифікує полігони міського атласу на 4 класи: зони першого пріоритету, зони першого пріоритету з проблемами для впровадження озеленення, зони другого пріоритету та зони третього пріоритету. Ці зони залежать від супутникових показників та їх тематичної класифікації на основі інформації про шкоду довкіллю. Дана

методологія була апробована на місті Ірпінь для якого було збудовано супутникові продукти аналізу міської поверхні та шару рекомендацій.

Також було розроблено сервіс моніторингу пожеж на основі супутникових продуктів Sentinel-3. Даний сервіс ґрунтується на модифікації алгоритму пошуку активного вогню розробленого для сенсорів супутникової місії MODIS. Він також містить аналіз температури земної поверхні, карт земного покриву та карт ґрунтів України, що дозволяє проводити класифікацію термальних аномалій як: торф'яна пожежа, лісова пожежа та паління стерні.

ВИСНОВКИ

В проєкті ERA-Planet SMURBS розв'язуються задачі аналізу та моніторингу розумних міст за допомогою супутникових даних. Температурні карти є важливим показником стану міста і може вказувати на проблемні зони та стихійні лиха такі як міські термальні острови та пожежі.

Основною задачею розв'язаною в даній дисертації є створення нового методу збільшення просторового розрізнення супутникових карт температури земного покриву Sentinel-3 на основі методів глибинного навчання та злиття мультиспектральних даних. Ця методологія ґрунтується на використанні рекурентної нейронної мережі LSTM з архітектурою "Всі в Один" яка бере на вхід часовий ряд супутникових продуктів NDVI та LST високого та низького розрізнення та будує карту температури земної поверхні для будь-якого дня для якого наявні ці дані. Особливістю методу є формування входу для поточного дня моделювання на основі диференціалу функції залежності температури земної поверхні від часу. Таким чином можна збільшити просторове розрізнення температурних карт Sentinel-3 від 1 км до 30 метрів, при цьому збільшивши часове покриття міста Київ супутниковими картами LST високого просторового розрізнення у двічі. Також в ході розв'язання даної задачі було розроблено нову методологію відновлення хмар на основі регресії випадкових лісів, що працює в режимі пошуку аномальних пікселів, які є дефектами знімку спричинені хмарами та тінями і моделювання безхмарної карти вегетаційного індексу NDVI та LST. Ця методологія збільшує якість вхідних даних які використовуються для збільшення просторового розрізнення супутникових продуктів Sentinel-3 та відповідно підвищує точність збільшення просторового розрізнення.

В даній дисертації було проведено прикладні дослідження для міста Рівне, Ірпінь та Київ на основі різнорідних супутникових даних. При

цьому карти температури земної поверхні є центральним показником, який дозволяє проводити аналіз планування міста розробку програми озеленення. Результатами цих досліджень є доступними у вигляді інтерактивних векторних карт опублікованих на державному геопорталі відкритих даних Управління містобудування та архітектури виконавчого комітету Рівненської міської ради а також на ресурсах Інституту Космічних Досліджень НАНУ та ДКАУ. На основі аналізу окремих кварталів міста Рівне із сегментацією по типу міського атласу відбувається розробка Рівненської програми озеленення міста ГО "Сад Історій Рівного". Також було розроблено методологію формування рекомендацій для озеленення міста на основі практик урбанізму інтерпретованих за допомогою супутникових даних. Використовуючи дану методологію було створено шар рекомендацій для міста Ірпінь. Також було реалізовано методологію виявлення пожеж на основі супутникових даних Sentinel-3 та поєднавши отримані карти пожеж за 2019-ий рік, карту ґрунтів та карту класифікації земного покриву було проведено аналіз пожеж в Україні. Виконання цих прикладних задач та розробка сервісів розумних міст проєкту SMURBS на основі відкритих супутникових даних викликали потребу у збільшенні просторового розрізнення карт температури земного розрізнення отриманих за допомогою супутникової місії Sentinel-3.

ПЕРЕЛІК ПОСИЛАНЬ

1. Sentinel-hub EO Browser: <https://www.sentinel-hub.com/explore/eobrowser>
2. Google Earth Engine Platform: <https://earthengine.google.com/>
3. Exploring Google earth engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping / [A. Shelestov, M. Lavreniuk, N. Kussul та ін.]. // Frontiers in Earth Science. — 2017. — №5.
4. Land Cover and Land Use Monitoring Based on Satellite Data within World Bank Project / [N. Kussul, A. Shelestov, M. Lavreniuk та ін.]. // 2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT). — 2019. — С. 127–130.
5. Lavreniuk M. Огляд методів машинного навчання для класифікації великих обсягів супутникових даних / M. Lavreniuk, A. Novikov. // Системні дослідження та інформаційні технології. — 2018. — №1. — С. 52–71.
6. A workflow for sustainable development goals indicators assessment based on high-resolution satellite data / [N. Kussul, M. Lavreniuk, A. Kolotii та ін.]. // International Journal of Digital Earth. — 2020. — №13. — С. 309–321.
7. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world / [P. Defourny, S. Bontemps, N. Bellemans та ін.]. // Remote sensing of environment. — 2019. — №221. — С. 551–568.
8. Satellite data reveal cropland losses in South-Eastern Ukraine under military conflict / [S. Skakun, C. Justice, N. Kussul та ін.]. // Frontiers in Earth Science. — 2019. — №7. — С. 305.
9. ASSESSMENT OF LAND CONSUMPTION FOR SDG INDICATOR 11.3.1 USING GLOBAL AND LOCAL BUILT-UP AREA

MAPS / [A. Shelestov, N. Kussul, B. Yailymov та ін.]. // IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium. — 2020.

10. Spatial Resolution ScienceDirect: <https://www.sciencedirect.com/topics/earth-and-planetary-sciences/spatial-resolution>

11. Chmiel J. Orthorectification and geometric quality assessment of very high spatial resolution satellite imagery for Common Agricultural Policy purposes / J. Chmiel, S. Kay, P. Spruyt. // In Proceedings of XXth ISPRS congress. — 2004. — C. 12–23.

12. Efficiency assessment of multitemporal C-band Radarsat-2 intensity and Landsat-8 surface reflectance satellite imagery for crop classification in Ukraine / [S. Skakun, N. Kussul, A. Shelestov та ін.]. // IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. — 2015. — №9. — C. 3712–3719.

13. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity / [F. Waldner, D. De Aballeyra, S. Verón та ін.]. // International Journal of Remote Sensing. — 2016. — №37. — C. 3196–3231.

14. Rahaman K. Pan-sharpening of Landsat-8 images and its application in calculating vegetation greenness and canopy water contents / K. Rahaman, Q. Hassan, M. Ahmed. // ISPRS International Journal of Geo-Information. — 2017. — №6. — C. 168.

15. Johnson B. An ensemble pansharpening approach for finer-scale mapping of sugarcane with Landsat 8 imagery. / B. Johnson, H. Scheyvens, B. Shivakoti. // International Journal of Applied Earth Observation and Geoinformation. — 2014. — №33. — C. 218–225.

16. Yuan, Qiangqiang, et al. "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11.3 (2018): 978-989.

17. Kussul, Nataliia, et al. "Assessment of Sustainable Development Goals Achieving with Use of NEXUS Approach in the Framework of

GEOEssential ERA-PLANET Project."XVIII International Conference on Data Science and Intelligent Analysis of Information. Springer, Cham, 2018.

18. Developing food, water and energy nexus workflows / [I. McCallum, C. Montzka, B. Bayat та ін.]. // International Journal of Digital Earth. – 2020. — №13. — С. 299–308.

19. Copernicus Land Monitoring Service Urban Atlas: www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-urban-atlas

20. CORINE Land Cover nomenclature conversion to Land Cover Classification system: https://land.copernicus.eu/eagle/files/eagle-related-projects/pt_clc-conversion-to-fao-lccs3_dec2010

21. Smart City Services for Kiev City Within ERA-PLANET SMURBS Project / [A. Shelestov, A. Kolotii, M. Lavreniuk та ін.]. // 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO). — 2019. — С. 784–788.

22. Urban Atlas for Ukrainian cities on the basis of high-resolution satellite data / [N. Kussul, A. Shelestov, B. Yailymov та ін.]. // Space Science and Technology. — 2019. — №25.

23. Essential variables for air quality estimation / [A. Shelestov, A. Kolotii, T. Borisova та ін.]. // International Journal of Digital Earth. — 2019.

24. Roadside collection of training data for cropland mapping is viable when environmental and management gradients are surveyed / [F. Waldner, N. Bellemans, Z. Hochman та ін.]. // International Journal of Applied Earth Observation and Geoinformation. — 2019. — №80. — С. 82–93.

25. Parcel based classification for agricultural mapping and monitoring using multi-temporal satellite image sequences / [N. Kussul, G. Lemoine, J. Gallego та ін.]. // 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). — 2015. — С. 165–168.

26. J. Jiménez-Muñoz, et al. "Land surface temperature retrieval methods from Landsat-8 thermal infrared sensor data."IEEE Geoscience and Remote Sensing Letters 11.10 (2014): 1840-1843.

27. J. Sobrino, et al. "Single-channel and two-channel methods for land

surface temperature retrieval from DAIS data and its application to the Barrax site."International Journal of Remote Sensing 25.1 (2004): 215-230.

28. Du, Chen, et al. "A practical split-window algorithm for estimating land surface temperature from Landsat 8 data."Remote Sensing 7.1 (2015): 647-665.

29. U. Avdan, and G. Jovanovska ., "Algorithm for automated mapping of land surface temperature using LANDSAT 8 satellite data." Journal of Sensors, 2016.

30. F. Wang, et al., "An improved mono-window algorithm for land surface temperature retrieval from Landsat 8 thermal infrared sensor data." Remote sensing, vol 7(4), pp. 4268-4289, 2015.

31. J. Jiménez-Muñoz, et al. "Improved land surface emissivities over agricultural areas using ASTER NDVI." Remote Sensing of Environment, vol 103(4), pp. 474-487, 2006.

32. Wan, Zhengming. "New refinements and validation of the MODIS land-surface temperature/emissivity products."Remote sensing of Environment 112.1 (2008): 59-74.

33. Lin, Shang-Hung, Sun-Yuan Kung, and Long-Ji Lin. "Face recognition/detection by probabilistic decision-based neural network."IEEE transactions on neural networks 8.1 (1997): 114-132.

34. Zangeneh, Erfan, Mohammad Rahmati, and Yalda Mohsenzadeh. "Low resolution face recognition using a two-branch deep convolutional neural network architecture."Expert Systems with Applications 139 (2020): 112854.

35. Lenka, Rasmita, et al. "Bio-Medical Image Processing: Medical Image Analysis for Malaria With Deep Learning."Examining Fractal Image Processing and Analysis. IGI Global, 2020. 158-169.

36. Deep learning classification of land cover and crop types using remote sensing data / N.Kussul, M. Lavreniuk, S. Skakun, A. Shelestov. // IEEE Geoscience and Remote Sensing Letters. — 2017. — №14. — С. 778–782.

37. Cloud Approach to Automated Crop Classification Using Sentinel-1 Imagery / [A. Shelestov, M. Lavreniuk, V. Vasiliev та ін.]. // IEEE Transactions

on Big Data. — 2019.

38. Fieuzal, Rémy, C. Marais Sicre, and Frédéric Baup. "Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks." *International journal of applied earth observation and geoinformation* 57 (2017): 14-23.

39. Skakun, S., et al. "The Use of Landsat 8 and Sentinel-2 Data and Meteorological Observations for Winter Wheat Yield Assessment." *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019.

40. Zhang, Qiang, et al. "Cloud and Shadow Removal for Sentinel-2 by Progressively Spatiotemporal Patch Group Learning." *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 2019.

41. Lavreniuk, Mykola, et al. "Use of Modified BM3D Filter and CNN Classifier for SAR Data to Improve Crop Classification Accuracy." *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*. IEEE, 2019.

42. Lavreniuk, M. (2019, July). Crop Classification Maps Filtration Based on Deep Learning Methods. In *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)* (pp. 1081-1085). IEEE.

43. Mikolov, Tomáš, et al. "Recurrent neural network based language model." *Eleventh annual conference of the international speech communication association*. 2010.

44. Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "On the difficulty of training recurrent neural networks." *International conference on machine learning*. 2013.

45. Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.

46. Gers, Felix A., Jürgen Schmidhuber, and Fred Cummins. "Learning to forget: Continual prediction with LSTM." (1999): 850-855.

47. Rußwurm, Marc, and M. Körner. "Multi-temporal land cover classification with long short-term memory neural networks." *The International*

Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 42 (2017): 551.

48. Copernicus Open Access Hub : <https://scihub.copernicus.eu/>

49. Sen2Core: <https://step.esa.int/main/third-party-plugins-2/sen2cor/>

50. Zhu, Zhe, Shixiong Wang, and Curtis E. Woodcock. "Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images." *Remote Sensing of Environment* 159 (2015): 269–277.

51. Claverie, Martin, et al. "The Harmonized Landsat and Sentinel-2 surface reflectance data set." *Remote sensing of environment* 219 (2018): 145–161.

52. Sentinelhub-py Python library: <https://github.com/sentinel-hub/sentinelhub-py>

53. SNAP : <https://step.esa.int/main/toolboxes/snap/>

54. Li, Wei, et al. "Gross and net land cover changes in the main plant functional types derived from the annual ESA CCI land cover maps (1992–2015)." (2018).

55. Use of land cover maps as indicators for achieving sustainable development goals / L.Shumilo, A. Kolotii, M. Lavreniuk, B. Yailymov. // IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. — 2018. — C. 830–833.

56. Transparent Land Governance in Ukraine within World Bank Program / [N. Kussul, A. Shelestov, M. Lavreniuk та ін.]. // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON). — 2019. — C. 1077–1080.

57. Nexus Approach for Calculating SDG Indicator 2.4. 1 Using Remote Sensing and Biophysical Modeling / N.Kussul, M. Lavreniuk, L. Shumilo, A. Kolotii. // IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. — 2019. — C. 6425–6428.

58. Lavreniuk, M., Shelestov, A., Kolotii, A., Vasiliev, V., Shumilo, L., Kussul, N. (2018, April). Automated System for Crop Mapping in Amazon

Web Services based on Sentinel Data. In EGU General Assembly Conference Abstracts (Vol. 20, p. 16629).

59. Sentinel-2 for agriculture national demonstration in Ukraine: results and further steps / [N. Kussul, A. Kolotii, A. Shelestov та ін.]. // 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). — 2017. — С. 5842–5845.

60. Blanco Abellán, Mónica, and Marta Ginovart Gisbert. "Introducing the normal distribution by following a teaching approach inspired by history: an example for classroom implementation in engineering education." Sixth Conference of European Research in Mathematical Education. 2009.

61. Lilliefors, Hubert W. "On the Kolmogorov-Smirnov test for normality with mean and variance unknown." Journal of the American statistical Association 62.318 (1967): 399-402.

62. Razali, Nornadiah Mohd, and Yap Bee Wah. "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests." Journal of statistical modeling and analytics 2.1 (2011): 21-33.

63. Rivne City Land Cover and Land Surface Temperature Analysis Using Remote Sensing Data / [L. Shumilo, B. Yailymov, N. Kussul та ін.]. // 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO). — 2019. — С. 813–816.

64. Land Surface Temperature Analysis and Trends for Rivne Using Remote Sensing Data / [L. Shumilo, B. Yailymov, M. Lavreniuk та ін.]. // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON). — 2019. — С. 1107–1111.

65. ГО "Сад Історій Півного": <https://treestory.in.ua/>

66. Генеральний План міста Рівне: <https://geo.rv.ua/ua/map/urban-plan#map=12//50.6191843//26.2302382&&layer=1664656129424098568-1,100//1758623072589448913-1,100>

67. Giglio, Louis, et al. "An enhanced contextual fire detection algorithm for MODIS." Remote sensing of environment 87.2-3 (2003): 273-282.

68. Land degradation estimation from global and national satellite based

datasets within UN program / [N. Kussul, A. Kolotii, A. Shelestov та ін.].
 // 2017 9th IEEE International Conference on Intelligent Data Acquisition
 and Advanced Computing Systems: Technology and Applications (IDAACS).
 — 2017. — С. 383–386.

69. Air Quality Estimation Using Satellite and In-situ Data for Kyiv
 City within ERA-PLANET Project / Y.Korsunska, L. Shumilo, A. Kolotii, A.
 Shelestov. // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer
 Engineering (UKRCON). — 2019. — С. 1032–1036.

70. Shelestov, Andrii, et al. "Indoor and outdoor air quality monitoring
 on the base of intelligent sensors for smart city." XVIII International Conference
 on Data Science and Intelligent Analysis of Information. Springer, Cham, 2018.